2025

The Impact of Data-Driven Learning of Grammar on the Development of English as a Foreign Language Students' Grammar Skills

(A Case of the University of Georgia)

Nino Tsulaia

Publication Partner:

IJSRP Inc.

The Impact of Data-Driven Learning of Grammar on the Development of English as a Foreign Language Students' Grammar Skills

(A Case of the University of Georgia)

Nino Tsulaia

Publishing Partner:
IJSRP Inc.
www.ijsrp.org

ISSN 2250-3153
9 772250 315302

Preface

Grammar is considered the backbone of a language, and many scholars emphasize the importance of grammar instruction in foreign language teaching. As a fundamental micro-skill, grammar plays a crucial role in developing and utilizing the four language skills. Proficiency in grammar contributes to overall language proficiency, making it essential to teach it engagingly and effectively.

The traditional method of teaching grammar is known as deductive instruction. This approach is fundamentally teacher-centered and often relegates students to a passive role, turning them into mere recipients of information rather than active constructors of knowledge. Such a passive learning environment stifles creativity and engagement, preventing students from becoming fully immersed in the learning experience. Consequently, the deductive method can feel monotonous and uninspiring, resulting in diminished interactivity and, ultimately, student demotivation and disengagement.

Therefore, there is an urgent need for a change in the approach to grammar teaching. It is crucial to transition from a teacher-centered classroom environment to a student-centered one. In this new setting, rather than simply transmitting information, teachers will facilitate students in constructing their knowledge. This shift will foster social interactions and create engaging and meaningful learning experiences. Moreover, 21st-century education requires the integration of digital technologies in teaching and an emphasis on developing students' higher-order thinking skills. It is essential to promote autonomous learning, enabling students to become self-directed learners rather than relying solely on their teachers.

In response to challenges, Corpus-based or Data-Driven Learning (DDL) offers an innovative, technology-oriented method for teaching grammar that provides authentic language-focused learning opportunities. DDL, also known as corpus-aided discovery learning, is a student-centered, constructivist approach that encourages inductive learning. This method actively engages essential learning skills, including analytical and critical thinking, collaboration, and problem-solving.

This monograph introduces readers to Data-Driven Language Learning, covering its practices and advantages. It discusses studies related to Data-Driven Learning of grammar, ways of corpus application in grammar instruction, and the impact on students’ skills. This piece of work proposes a model for Data-Driven Learning of grammar, which was used in an experimental study. The monograph presents the findings of the experimental study conducted at the University of Georgia, aiming to measure the development of undergraduate students’ grammar skills through Data-Driven Learning. Additionally, the paper presents the results of a survey that

explored students' attitudes towards Data-Driven Learning of grammar in the experimental groups.

The presented work is intended for EFL (English as a Foreign Language) teachers seeking to move away from the traditional, tedious practice of grammar instruction and implement a modern, technology-based method that involves active, constructivist learning, authentic language teaching, and the development of skills necessary for the 21st-century learner.

Publication Partner:
International Journal of Scientific and Research Publications (ISSN: 2250-3153)

Copyright and Trademarks

Authors

Nino Tsulaia
Doctor of Education Sciences, assistant professor at Sokhumi State University
Tbilisi, Georgia

Table of Content

List of Figures

List of Tables

1. DATA-DRIVEN LANGUAGE LEARNING

Technological advances over the last century have led to the development of a groundbreaking trend in linguistics known as corpus linguistics. Although corpus linguistics originated in the 1950s, it has undergone significant evolution in recent decades. This field involves studying a language by analyzing language data that is collected in a corpus.

A corpus is defined by O'Keeffe, McCarthy, and Carter (2007) as "a collection of texts, written or spoken, which is stored on a computer" (p. 1). Similarly, Reppen (2010) refers to a corpus as "a large, principled collection of naturally occurring texts (written or spoken) stored electronically" (p. 23). A corpus serves as a powerful tool for the linguistic description and analysis of a language.

There are several types of corpora, including general and specialized corpora, written and spoken corpora, synchronic and diachronic corpora, monolingual and multilingual corpora, parallel and comparable corpora, annotated and unannotated (or raw) corpora, native speaker and learner corpora, as well as static and monitor corpora. A general corpus includes both oral and written texts, while a specialized corpus is genre-specific and focuses on a particular type of text. Synchronic corpora consist of texts from a specific period, typically the present time, whereas diachronic (or historical) corpora include texts from various historical periods, making them valuable for analyzing language development. Monolingual corpora contain texts in only one language, while multilingual corpora can be divided into two types: parallel and comparable. A parallel corpus consists of texts and their translations in different languages, whereas a comparable corpus contains similar texts in different languages, allowing for comparisons of language patterns in similar situations. An annotated corpus includes linguistic information (metadata) added to the texts, while an unannotated (or raw) corpus consists of texts without any additional linguistic information. A learner corpus is made up of texts produced by language learners. A static corpus remains fixed, with no further updates, whereas a monitor corpus is continuously updated, with recent texts added regularly.

Several distinguished and widely used web-accessible English corpora include the British National Corpus (BNC), the Corpus of Contemporary American English (COCA), the Brown Corpus, the Bank of English, the International Corpus of English (ICE), the Michigan Corpus of Academic Spoken English (MICASE), Cambridge Academic English, the English SKELL Corpus, EUROPARL, English Web, and the Corpus of Historical American English (COHA), among others. There are also web-based software tools that provide easy access to a multitude of diverse corpora. For example, Sketch Engine offers around 500 corpora in more than ninety languages. CQPweb, from Lancaster University, provides access to contemporary English corpora, historical English corpora, historical news text corpora, Shakespearean corpora, and more.

The development of corpus linguistics has had a significant impact on the teaching of foreign languages. In recent decades, the potential benefits of using corpora in language education have been increasingly recognized by scholars and educators. The use of corpora in language instruction began in the 1980s, and this approach is known as Data-Driven Learning (DDL). The term DDL was introduced by Tim Johns in 1990 (Johns, 1990). Johns, a British academic and professor of English as a Foreign Language at the University of Birmingham, anticipated a revolution in applied linguistics and language teaching through the use of corpora, highlighting their potential for enhancing language education.

Data-Driven Learning (DDL) is a form of Computer-Assisted Language Learning (CALL) that is beneficial for teaching first, second, or foreign languages, as reported by Al-Gamal and Ali (2019). DDL is also known as corpus-based learning, according to Geluso and Yamaguchi (2014). This approach can utilize both native and learner corpora.

Corpora offer significant support for the language learning process and are widely used in applied linguistics. Various corpus products include textbooks for vocabulary and grammar instruction, grammar reference books, bilingual dictionaries, dictionaries of idioms and proverbs, dictionaries of collocations, grammar dictionaries, and dictionaries of phrasal verbs, among others.

Corpora can be utilized for various pedagogical purposes. They enable teachers to design course books, grammar guides, usage manuals, and more. It's important to note that corpus analysis can influence the teaching curriculum and impact the design of syllabi and educational materials. Using corpus data, teachers have the opportunity to create syllabi tailored to observed language usage. This data allows educators to prioritize certain topics that occur

more frequently within the language system (De Jesus & Carrilo, 2014). Consequently, teaching materials can be developed based on the insights gained from the corpus.

Figure 1 below illustrates the capabilities of corpora for language teaching and corpus-based outputs.

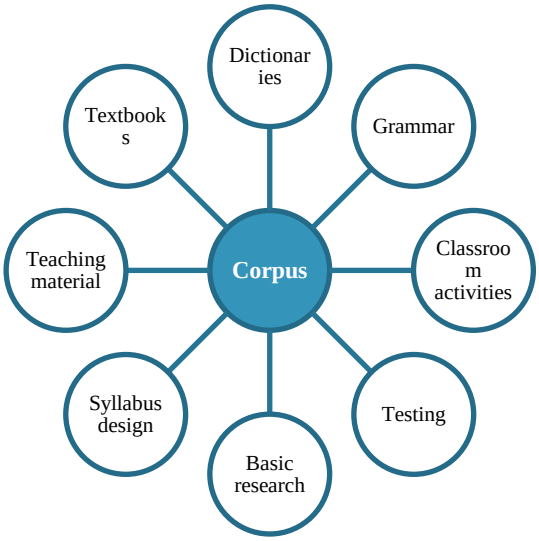

Figure 1. Using corpora in the context of language teaching (Source: Johansson, 2009, p. 40).

Scholars differentiate between two types of applications of corpora in the classroom: indirect and direct applications. This distinction leads to two forms of Data-Driven Learning (DDL): indirect DDL and direct DDL. Indirect DDL involves the teacher designing corpus-based materials, which students then use to identify and analyze language features. In contrast, direct DDL allows learners to interact directly with computers and corpus software, enabling them to explore authentic corpus data to investigate language patterns such as grammar structures, collocations, and frequency indicators. Indirect DDL is often considered a paper-based approach, while direct DDL is categorized as a computer-based or hands-on approach.

Marinov (2013) distinguishes between two approaches to using corpora in the classroom: the soft version and the hard version. The soft version, referred to by some scholars as indirect Data-Driven Learning, involves providing students with corpus-based handouts. In contrast, the hard version aligns with direct Data-Driven Learning, where students interact directly with a corpus. Sun and Hu (2020) associate the use of corpus-based reference materials - such as dictionaries, grammar books, and language learning coursebooks, which are based on corpus data - with the indirect approach to corpus-based learning. Forms of corpus exploitation in teaching and learning processes are summarized in Figure 2 below:

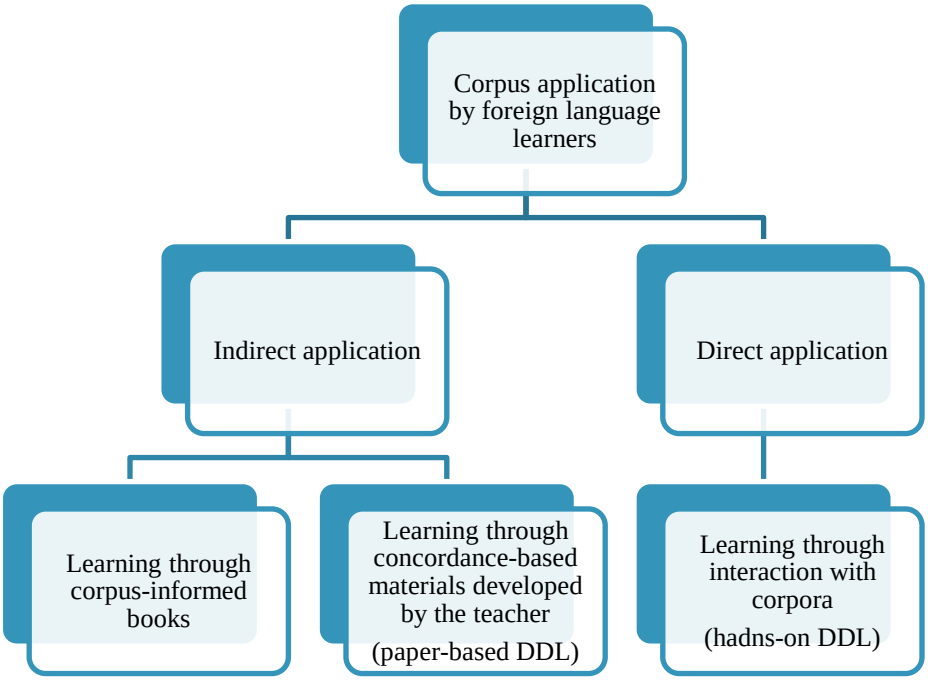

Figure 2. Types of corpus application by foreign language learners.

Concordance-based materials are particularly appealing to learners. A concordance is an essential component and a powerful tool for searching through countless examples of a specific word or phrase within a corpus. It presents target words in various contexts through concordance lines. A common format for these lines is known as keyword-in-context (KWIC), which displays the target word or phrase in the center, flanked by other words on the left and right. Teachers can edit the contents of the concordance when developing materials. Reppen (2010) notes that corpus-based tasks do not necessarily need to differ from the activities and exercises that students and teachers are accustomed to in conventional teaching. They can maintain a similar format to ordinary language-learning tasks. Likewise, Kazuko (2014) indicates that paper-based concordance lines can be introduced to learners in their original, unaltered form or revised, which may benefit beginner-level language learners.

Chujo, Anthony, Oghigian, and Uchibori (2012) classify Data-Driven Learning (DDL) into three categories: computer-based, paper-based, and combined methodologies. In computer-based DDL, students interact directly with corpora, use corpus tools, and work with the data displayed in a specific corpus. In contrast, paper-based DDL involves students working with concordance materials developed by the teacher. Combined DDL incorporates elements from both methodologies. According to Chujo et al. (2012), paper-based DDL is more time-efficient compared to the computer-based approach. However, paper-based DDL is more accessible, as students do not need to use computers. Furthermore, tasks in paper-based DDL are generally easier for students to understand (Chojo et al., 2012). Studies such as those by Boulton (2012) and Yoon and Jo (2014) have demonstrated that students engaged in paper-based DDL tend to achieve slightly higher scores than those who participate in computer-based DDL.

Language data provided by corpora is an invaluable resource for language learning. A corpus is a collection of a vast amount of words and phrases gathered from various written and audio materials, including books, newspapers, TV and radio broadcasts, and online content. This means that corpora encompass both formal and informal language. As noted by Gilquin and Granger (2010), “Raw corpora offer numerous possibilities for the exploration of language by learners” (p. 361). Specialized corpora are particularly beneficial for specific purposes. Reppen (2010) points out that, due to technological advancements, today's corpora are much larger than those from previous decades, although earlier corpora are still in use. The author suggests that a one-million-word general corpus is sufficient for observing language patterns and can be used effectively for teaching.

Data-Driven Learning is an approach that exposes learners to a large volume of authentic corpus data. Corpora provide easy and immediate access to a wealth of real-world language examples. Students become familiar with

instances from a wide array of authentic texts, making Data-Driven Learning a genuinely authentic language learning process. According to De Jesus and Carrilo (2014), a corpus is a valuable tool for teachers using conventional methodologies, as it allows them to reference authentic examples rather than relying solely on made-up sentences found in traditional textbooks.

The data from corpora are genuine, diverse, illustrative, and up-to-date. They showcase language as it is used in real contexts, reveal language trends, and offer more varied and valuable information than traditional dictionaries and reference grammars. Many scholars, such as Dazdarevic, Zoranic, and Fijuljanin (2015), prefer using corpora over dictionaries for understanding word meanings.

Data-Driven Learning significantly enhances students' language awareness. According to Dazdarevic, Zoranic, and Fijuljanin (2015), “corpus helps students get a broader view of language” (p. 5). Corpora enable learners to recognize the frequency of different language features and identify “the most common and frequent lexical choices” made by native speakers (Dash, 2011, p. 21). As Huang (2011) reports, “corpus data enrich our understanding of language use” (p. 482).

The Data-Driven Learning (DDL) method encourages active learning among students. In DDL, students engage deeply in examining various aspects of language. They infer the meanings of words and expressions while also observing their usage and frequency. Those who participate in Data-Driven Learning are often described as travelers, detectives, or even "Sherlock Holmes." According to Chirobocea (2017), the corpus-based approach to language teaching aims to "make linguistic detectives out of learners" (p. 365).

Data-Driven Learning (DDL) involves noticing language patterns within concordance lines. The Noticing Hypothesis, developed by Schmidt (1993), suggests that learners often do not acquire linguistic features unless they consciously notice them. In his work from 1990, Schmidt emphasizes the significance of consciousness and the consciousness-raising process in language learning (Schmidt, 1990). He identifies three dimensions of consciousness: awareness, intention, and knowledge. Within awareness, he distinguishes three levels: perception, noticing (focal awareness), and understanding. Schmidt argues that "noticing" is a crucial stage in transforming input into intake. The Noticing Hypothesis is a key theoretical basis for Data-Driven Learning.

Data-Driven Learning (DDL) is a constructivist teaching method that promotes self-directed learning. In this approach, students are encouraged to become self-regulated learners. By utilizing corpus-aided teaching, students gain an effective strategy for "learning how to learn." They “work independently or collaboratively to observe, analyze, and interpret patterns of language use” (Huang, 2008, p. 21).

Data-Driven Learning (DDL) transforms the roles of both teachers and students. The integration of a corpus into the teaching and learning process fosters a more student-centered environment, where students take on the role of researchers while teachers act as consultants.

Teaching with corpora enhances student autonomy, often leading to increased responsibility for their own learning. Students explore language features within large datasets. As noted by De Jesus and Carrillo (2014), students engaged in corpus-aided language learning "conduct their own research about language features" (p. 4).

A corpus is considered a valuable tool for deepening language learning. Corpora provide numerous opportunities for language acquisition, and Data-Driven Learning (DDL) is applied to various aspects of language, including vocabulary and grammar. DDL is particularly effective in teaching collocations and phraseology. Corpus-based materials enable learners to improve their linguistic skills, encompassing synonyms, antonyms, polysemy, homonymy, and more (Al-Gamal & Ali, 2019). A corpus serves as a resource for understanding collocation, colligation, and the prosody of target words. Collocation refers to the common combination of words, while colligation pertains to patterns of lexical and grammatical bundles. The context in which collocation and colligation occur is known as semantic prosody. Semantic prosody can convey three types of connotational value: positive, neutral, or negative (Mansoor & Salman, 2013). All these aspects can be effectively learned through the use of corpora. Corpora are also extensively utilized in teaching translation, facilitating error self-correction, and improving writing skills. Additionally, corpus data can be employed to create materials for other language skills, such as reading and speaking.

Data-Driven Learning, also known as discovery learning with corpora or corpus-aided discovery learning (CADL), involves exploratory tasks and activities. Huang (2008) distinguishes between guided and unguided approaches in corpus-aided discovery. Barabadi and Khajavi (2017) argue that Data-Driven Learning is no longer viewed as a purely discovery-based method, as it was initially perceived. Many scholars emphasize the advantages of teacher-

guided inquiry. They support the idea that using teacher-prepared corpus material is much more effective than simply exposing students to raw data from the corpus.

A corpus can be utilized in both deductive and inductive approaches to language teaching. In the deductive approach, corpus analysis helps teachers and students understand how language elements are actually used by native speakers. Conversely, in the inductive approach, corpus data allows students to make generalizations and deduce language rules (De Jesus & Carrilo, 2014).

Data-Driven Learning provides students with the opportunity to develop essential learning skills. It enhances both cognitive and metacognitive abilities. According to O'Sullivan (2007), a variety of cognitive skills are involved in the corpus-based learning process, including: predicting, observing, noticing, thinking, reasoning, analyzing, interpreting, reflecting, exploring, making inferences (inductively or deductively), focusing, guessing, comparing, differentiating, theorizing, hypothesizing, and verifying.

Moreover, Data-Driven Learning encourages the incorporation of problem-solving activities, which help foster critical thinking skills. Dash (2011) points out that corpora can be used to critically evaluate existing English Language Teaching (ELT) materials. Often, teaching textbooks contain fictional language that diverges significantly from the language used by native speakers in real-life situations.

According to Braun (2007), it's essential to shift our focus toward needs-driven corpora and materials. Today, teachers and students have the opportunity to create their own corpora based on their specific needs and objectives. Boulton (2016) argues that allowing learners to develop their own small corpora fosters "a sense of ownership and familiarity with the contents" (p. 124).

Data-Driven Learning encourages generativity, prompting students to actively use language. This approach aligns with students' interests and makes the discovery aspect of the Data-Driven Learning method engaging, which adds an element of fun to the learning process. Furthermore, it promotes interactivity within the classroom and enhances knowledge retention (Brown, 2017).

A substantial body of research demonstrates the advantages of corpus-based instruction over traditional teaching methods in various aspects of language education. Studies conducted by Barabadi and Khajavi (2017), Binkai (2012), Jalilifar, Mehrabi, and Mousavinia (2014), and Unaldi et al. (2013) have validated the effectiveness of the corpus-based approach in enhancing EFL learners' vocabulary acquisition. Additionally, the research by Elsherbini and Ali (2017) revealed a positive impact of corpus-based activities on EFL students' grammar and vocabulary knowledge. Numerous studies (e.g., Ashouri, Arjmandi & Rahimi, 2014; Chan & Liou, 2005; Çelik, 2011; Li, 2017; Vyatkina, 2016) have also demonstrated the benefits of Data-Driven Learning (DDL) on the collocation competence of EFL learners. Koosha and Jafarpour (2006) showed significant improvement in EFL learners' understanding of prepositional collocations through the DDL method. Furthermore, the research by Jafarpour, Hashemian, and Alipour (2013) confirmed the effectiveness of DDL in enhancing EFL learners' comprehension and production skills regarding collocations of synonyms. Ucar and Yükselir (2015) highlighted the positive effects of corpus-based teaching on EFL learners' knowledge of verb-noun collocations. Moreover, the study by Rahimi and Momeni (2012) indicated that corpus-based collocation instruction improved students' overall language proficiency. Kayaoglu (2013) provided evidence for the efficacy of using corpus techniques to teach synonyms to English language learners. Similarly, Boontam and Phoocharoensil (2018) affirmed the effectiveness of DDL in learning English prepositions. Çelik and Elkatmiş (2013) showcased the advantages of corpus-assisted language teaching in helping learners acquire punctuation marks. A considerable number of studies (e.g., Boontam & Phoocharoensil, 2018; Boulton, 2010; Elsherbini & Ali, 2017; Chujo, Utiyama & Miura, 2006; Geluso & Yamaguchi, 2014; Hadley, 2002; Smith, 2011; Sun & Hu, 2020; Takahashi & Fujiwara, 2016; Varley, 2009; Vyatkina, 2016; Yoon & Hirvela, 2004) indicate that students have positive attitudes towards Data-Driven Learning.

Data-Driven Learning (DDL) presents several challenges and disadvantages. One significant issue is the difficulty in applying corpora, which is often linked to concerns about language authenticity. According to Brown (2017), some teachers worry that corpus data might "confuse or overwhelm learners" (p. 2). Additionally, Gilquin and Granger (2010) point out that corpora can contain "non-standard forms, swear words, or literary phrases" (p. 367). Kazuko (2014) concurs with this view, noting that corpus data includes non-ordinary structures.

Another challenge is that inspecting and evaluating corpus data requires a considerable investment of time. A study by Elsherbini and Ali (2017) found that students who were taught using DDL faced difficulties in analyzing the data because it demanded significant time and effort from them.

While DDL is often considered suitable primarily for advanced learners, many scholars argue that it can benefit students at all levels of language proficiency. For instance, Al-Gamal and Ali (2019) assert that the corpus-based method is advantageous for learners across the proficiency spectrum. Furthermore, several studies (e.g., Boulton, 2008; Boulton, 2009; Chujo, Utiyama & Miura, 2006; Takahashi & Fujiwara, 2016; Yunus, 2014) have demonstrated the effectiveness of DDL with elementary-level students as well. To address these challenges, especially for learners with lower language skills, researchers recommend editing paper-based concordance lines and removing confusing examples from them.

Some scholars (e.g., Flowerdew, 2009; Gilquin & Granger, 2010) argue that Data-Driven Learning (DDL) may not accommodate all students' learning styles. This is largely due to its emphasis on inductive learning, which might not be suitable or desirable for every type of learner. Another concern is the challenge of mastering the software involved in DDL. Vyatkina (2020) notes that “most corpora have been developed by corpus linguists for corpus linguists, with interfaces that are too technical and difficult for non-specialist users, such as language teachers and learners” (p. 363). Therefore, both teachers and learners must develop corpus literacy and receive training in this area.

Additionally, a significant barrier to incorporating corpora into teaching is that many corpora are not freely accessible to students. Furthermore, preparing corpus-based materials for teaching can be extremely time-consuming. Gilquin and Granger (2010) elaborate that it takes a considerable amount of time to create concordance-based materials, to familiarize students with corpus tools, and to have them engage in Data-Driven Learning activities. According to Huang (2018), extensive implementation of corpus-assisted learning requires considerable time and effort from teachers, particularly during the initial phase.

Data-Driven Learning (DDL) is not widely used as a teaching method around the world, and corpora are not extensively utilized as teaching aids in classrooms. Scholars such as Frankenberg-Garcia (2012), Jalilifar, Mehrabi, and Mousavinia (2014), Römer (2010), and Qiong (2017) emphasize the importance of raising awareness about corpora and corpus-based teaching, as many teachers are unfamiliar with what a corpus is and the benefits it can offer in language pedagogy.

There are conferences held globally that focus on the potential of various types of corpora and their effectiveness in language acquisition. These conferences bring together teachers, methodologists, researchers, and software engineers. Additionally, numerous studies have been published to demonstrate the effectiveness of corpora and promote their widespread use in language teaching. Römer (2010) advocates for the creation of a “DDL-friendly environment” that encourages both direct and indirect use of corpora in language instruction.

In summary, despite certain challenges, Data-Driven Learning is an effective constructivist approach and a form of Computer-Assisted Language Learning that is well-suited to meet the needs and interests of students in today's educational landscape. A corpus provides unique opportunities for language acquisition. DDL is a revolutionary method of foreign language teaching that fosters cognitive competencies and valuable lifelong skills in students, including analytical skills, problem-solving abilities, decision-making skills, critical thinking skills, and collaboration skills, among others.

The field of corpus linguistics is actively developing within Georgian higher education. Since the late 1990s, several Georgian corpora have emerged in the study of linguistics. These include:

- TITUS-ARMAZI (Alternative Resources, Materials, Applications, and Zipped Information on the TITUS University platform) (1999-2002);
- ECLinG (Endangered Caucasian Languages in Georgia) (2002-2005);
- SSGG (Die soziolinguistische Situation im gegenwärtigen Georgien/The Sociolinguistic Situation of Present-Day Georgia) (2005-2007).

Building upon these resources, the Georgian National Corpus was later developed. Additional projects that contributed to the establishment of the Georgian National Corpus include:

- Georgian Dialect Corpus (created in 2009); Georgian Scientific Metalanguage Corpus (created in 2016);
- Georgian Political Corpus (created in 2016).

The current Georgian National Corpus comprises several subcorpora, which include Old Georgian, Middle Georgian, Modern Georgian, the Georgian Reference Corpus, the Georgian Dialect Corpus, Political Texts, Old and Middle Georgian Law Texts, as well as corpora for Megrelian and Svan languages (GNC, 2022; Kenchianchvili, 2020).

Various corpora have been created through different projects at Georgian higher education institutions, such as:

- The Georgian Language Corpus at the Institute of Linguistic Studies of Ilia State University (since 2009);
- The Epigraphic Corpus of Georgia at the Institute of Linguistic Studies of Ilia State University (since 2015);
- The Georgian Language Linguistic Corpus at Georgian Technical University (since 2013);
- The English-Georgian Parallel Corpus of Scientific Texts at the Lexicographic Center at Tbilisi State University (since 2013).

Additionally, there are two web corpora:

- KaWaC (Kartvelian Web as a Corpus);
- OpenSubtitles Parallel Corpora (Georgian is included in the list of languages) available on the Sketch Engine corpora server (Kenchianchvili, 2020; Tandashvili & Purtskhvanadze, 2014).

Textbooks on corpus linguistics have been published in the Georgian language since 2014. The course titled "Introduction to Corpus Linguistics" has been offered at Tbilisi State University since 2008 and at Akhaltsikhe State University since 2012. Ilia State University has been conducting a doctoral program in Digital Humanities since 2017, with one of its focuses on Computer and Corpus Linguistics. Additionally, Batumi State University has been running a minor BA program in Digital Humanities since 2016. Summer and winter schools in corpus linguistics have been organized in collaboration with various universities since 2012, and students from different institutions participate in these seasonal schools, contributing to projects aimed at creating corpora (Kenchiashvili, 2020).

However, corpora have not yet been utilized for foreign language teaching purposes at the general education or higher education levels in Georgia, and as a result, the application of corpora in language pedagogy remains an unexplored area.

2. STUDIES RELATED TO DATA-DRIVEN LEARNING OF GRAMMAR

The Data-Driven Learning (DDL) method is commonly used in teaching foreign language grammar. Both hands-on and paper-based DDL approaches are actively implemented in grammar instruction. However, scholars like Anani Sarab and Kardoust (2014) suggest that the direct use of corpora in the classroom is relatively uncommon.

In the DDL approach to grammar, learners examine target items within concordance lines, which can be accessed through printouts or software concordancers. These concordance lines consist of real language data, allowing students to investigate the data, make discoveries, analyze it, and infer rules based on their thorough examination. Boontam and Phoocharoensil (2018) state that the “DDL approach is an inductive learning strategy” (p. 127). In this framework, the teacher acts as a guide, facilitating the students' deduction process (Dazdarevic, Zoranic, & Fijuljanin, 2015). According to Koosha and Jafarpour (2006), the “DDL approach suggests that grammar learning should consist largely of consciousness-raising activities rather than the teaching of rules” (p. 196). Consequently, DDL students do not learn grammar rules explicitly.

Johns (1991), the pioneer of Data-Driven Learning (DDL), proposed a three-step, concordance-based learning method (Figure 3). He highlights the significant benefits of DDL, stating that it helps learners "develop the ability to see patterns in the target language and to form generalizations to account for those patterns" (p. 2). This approach transforms the role of the teacher, who becomes a "coordinator in student-initiated research" (Johns, 1991, p. 3). DDL revolutionizes grammar instruction by enabling learners to discover grammatical rules within an authentic context, enhancing their grammatical awareness.

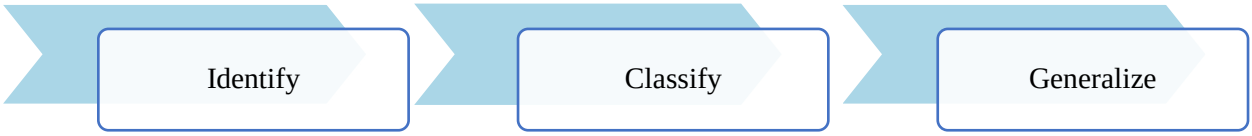

Figure 3. The Data-Driven Learning principle (based on Johns, 1991, p. 4).

Corpora allow for the study of grammar from a descriptive rather than a prescriptive perspective, focusing on how language is actually used in reality (Jones & Waller, 2015). Biber and Reppen (2002) note that corpus data often reveals significant differences between the way native speakers use language and what students learn from traditional grammar textbooks.

Carter and McCarthy (1995) developed an inductive, corpus-based approach to grammar instruction, which consists of three stages known as the three "I's": illustration, interaction, and induction (Figure 4). The illustration stage involves examining the data; the interaction stage includes discussions, considerations, exchanging views, and exploration; and the final stage of induction entails drawing conclusions and formulating rules for specific linguistic elements.

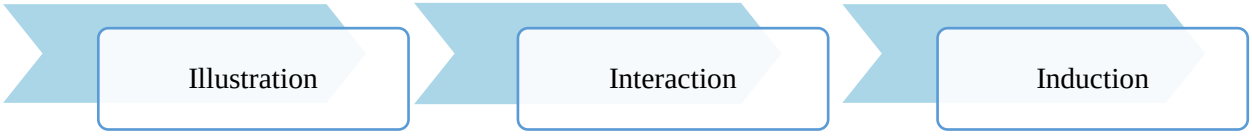

Figure 4. Inductive corpus-based learning of grammar (based on Carter and McCarthy, 1995, p. 14).

Flowerdew (2009), drawing on Carter and McCarthy's (1995) approach to learning, developed a new paradigm for corpus-based learning of lexicogrammar, specifically focusing on collocations, as summarized in Figure 5. He introduced a non-compulsory phase of intervention that precedes the final phase of induction. The intervention phase involves the teacher providing clues, indications, or directions to learners. Flowerdew explains that inductive learning often encounters challenges related to different learner types or the interpretation of corpus data. To address these issues, he recommends incorporating the intervention phase to guide students toward the correct conclusions.

Figure 5. A paradigm for corpus-based learning of lexicogrammar (based on Flowerdew, 2009, p. 407).

Chujo and Oghigian (2007) conducted a study with twenty beginning-level English learners at a Japanese university, utilizing a blended approach that combined Data-Driven Learning (DDL) for grammar with Computer-Assisted Language Learning (CALL) for vocabulary. The research spanned two semesters. During the first semester, students engaged in DDL activities in pairs, while in the second semester, they completed tasks individually. The post-test results indicated significant improvements, with students demonstrating enhanced overall English proficiency and a better understanding of the targeted grammar points, such as the number of nouns, word classes, verb forms, and adverbs. Questionnaire responses revealed that most students enjoyed the DDL/CALL course, finding it both beneficial and enjoyable, with a strong inclination to continue their studies. Additionally, students evaluated the DDL method and parallel concordance activities positively. The majority found these activities to be meaningful, valuable, clear, and easy to use. Most students also considered DDL to be effective for learning grammar, vocabulary, and retention.

Chujo and Oghigian (2008) conducted a study to explore the effects of the Data-Driven Learning (DDL) approach on beginner-level English as a Foreign Language (EFL) learners' acquisition of noun phrases and verb phrases. This research took place at a university in Japan and spanned one academic year. The focus was on specific grammatical features: noun phrases, which included various word classes, countable and uncountable nouns, and a range of noun phrase structures; and verb phrases, which encompassed tenses, passive voice constructions, clauses, infinitives, gerunds, and more. At the end of the academic year, the students demonstrated significant improvement on the post-test. Additionally, they completed a questionnaire evaluating the DDL worksheets, the presentation of grammar,

follow-up activities, the feedback received from the teacher, and their overall experience with the DDL course. The students reported that Data-Driven Learning was highly effective, noting that the feedback they received was particularly valuable, as it guided them to the correct conclusions.

Chujo and Oghigian (2008) outlined a procedure for Data-Driven Learning (DDL) of grammar, which can be summarized as follows (Figure 6). The first stage involves the teacher presenting DDL materials, after which learners observe grammatical patterns within those materials and formulate hypotheses based on their observations. In the second stage, students either verify or revise their hypotheses through the teacher's explanations and clarifications. The third stage involves learners practicing and testing their hypotheses through follow-up exercises. Finally, in the fourth stage, students apply their hypotheses in production activities, meaning they create language patterns. Throughout the third and fourth stages, the teacher provides feedback to the students.

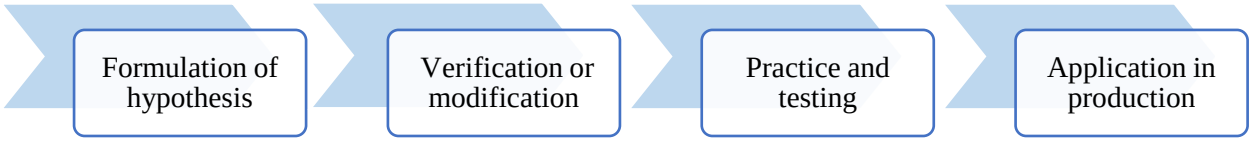

Figure 6. A DDL approach to learning grammar (based on Chujo and Oghigian, 2008, pp. 66-67).

Chujo, Anthony, Oghigian, and Yokota (2013) investigated the effectiveness of Data-Driven Learning (DDL) for low (remedial) English proficiency students at a Japanese university. The study involved twenty-two students and spanned fifteen weeks. The DDL syllabus covered various grammar points, including nouns, adjectives, modals, pronouns, conjunctions, and indirect questions. The experiment followed a four-step procedure outlined by Chujo and Oghigian (2008). The first step involved forming hypotheses through inductive tasks using hands-on corpus data. This was followed by explicit explanations from the teacher. The final steps included follow-up exercises, production tasks, and feedback from the teacher. During the experiment, parallel corpora were utilized. The post-test results indicated significant improvements in the students' understanding of the targeted grammar points. Both the teacher and the students reported positive attitudes toward the integration of parallel corpora into the teaching and learning process. Most students found the ability to translate into their native language using the parallel corpus to be very helpful and necessary.

Chujo, Kobayashi, Mizumoto, and Oghigian (2016) conducted a three-year study involving freshman university students in Japan who were beginner EFL (English as a Foreign Language) learners. The study focused on a DDL (Data Driven Learning) syllabus aimed at teaching different types of noun phrases. The instructional approach followed the four-stage DDL teaching procedure outlined by Chujo and Oghigian (2008). In the first year, the students in the first group used only one concordance tool, WebParaNews. The second group of students in the second year utilized two concordance tools: WebParaNews and LagoWordProfiler for ParaNews. In the third year, the students in the third group also used the same two tools, WebParaNews and LagoWordProfiler for ParaNews. All three groups made significant progress; however, the group that used two tools achieved notably higher results compared to the single tool groups. Furthermore, participants in the double tool group reported more positive evaluations on the attitude questionnaire.

Boulton (2009) conducted a study with first-year college students in France who were studying English. The focus was on ten linking adverbials: *but, however, actually, in fact, anyway, whereas, on the other hand, besides, nonetheless, and on the contrary.* The students' English proficiency levels ranged from elementary to intermediate. Four groups were formed for the experiment. Two groups learned the linking adverbials using corpus information: one group engaged with short context (SC) examples, while the other group used key-in-context (KW) examples. The other two groups learned the adverbials through traditional pedagogical methods: one group utilized a bilingual dictionary (BD), and the other focused on grammar usage (GU). The results of the study indicated that corpus-based resources were more effective than traditional language learning methods. Additionally, Data-Driven Learning proved to be beneficial for lower-level learners as well.

Rapti (2010) conducted a study to investigate the effects of Data-Driven Learning (DDL) on school students' grammar acquisition and motivation in Greece. The participants were at the B1 level, and the treatment lasted for five months. In this study, DDL involved providing concordance-based grammar materials and tasks tailored to the students. The grammar topics covered included tenses, conditionals, passive voice, reported speech, gerunds, infinitives, causative forms, modal verbs, and more. Overall, students in the experimental group demonstrated

significant improvements in their test results. Their attitudes towards the DDL materials, activities, and tasks, as well as their motivation levels, were assessed through a questionnaire and interviews. Most students reported a positive experience with DDL for grammar, although they expressed a preference for future grammar learning to include traditional grammar books as well. The learners faced challenges related to working with concordances, language complexity, and incomplete context (KWIC format). The author suggests combining inductive and deductive approaches for the effective implementation of the DDL method.

Sahillioglu, Şahinkayası, and Şahinkayası (2011) examined the effectiveness of two corpus-based approaches on Turkish students' learning of English grammar. The study involved 77 low-level English learners and lasted for three weeks. Two different teaching approaches were implemented during this period. One group was instructed using AntConc, a free software that allows users to build a small or specialized corpus from selected materials. For this group, books deemed appropriate for beginner students were chosen to construct their corpus. The other group was taught using the extensive, web-based BYU-BNC (British National Corpus, managed by Brigham Young University in Provo, Utah). The objective of the study was to compare the effectiveness of these two teaching methods. Both groups utilized a concordance tool as part of the Data-Driven Learning process, focusing on the teaching of tenses and modals. The results of an achievement test indicated that both corpus tools improved students' understanding of grammatical structures. However, there was no significant difference in the gains between the two groups; thus, both approaches proved to be equally effective. Interestingly, the part-of-speech (POS) query feature in AntConc was found to be more effective for constructing grammatical structures than the BYU-BNC. Despite the improvements in grammatical understanding, neither tool enhanced the learners' self-confidence in using target grammar features in their writing. Furthermore, the attitudes of both groups toward the use of their respective corpus tools to learn English grammar declined after the treatment period. The researchers noted that the students struggled to deduce meaning from the concordance lines. They suggested that shifting the teaching approach to a more deductive mode could enhance student engagement. Teachers faced time constraints in training and teaching students, which impacted the learning process. Many students felt uncertain about handling authentic language data. In conclusion, the researcher recommends using a combination of corpus-based and deductive teaching methods for low-level students and suggests conducting a longer-term study. The results from the attitude questionnaire indicated that students found Data-Driven Learning (DDL) to be useful for vocabulary acquisition. They appreciated DDL for its benefits in understanding the meanings of words, recognizing collocations, and learning word usage. Most students believed that their writing skills had improved as a result of DDL. Additionally, the majority of students felt that DDL was not time-consuming, and they reported having little difficulty with unfamiliar words in concordance lines or in inferring usage rules for words.

Girgin (2011) examined the impact of corpus-based activities on grammar acquisition among lower-level EFL university students in Turkey. The targeted grammar points included relative clauses, conditionals, the passive voice, indirect questions, and question tags. The treatment lasted for three weeks. Girgin's corpus-based activities involved presenting concordance lines followed by leading questions and gap-filling exercises. The results indicated that while both the experimental and control groups showed similar achievements, corpus-based activities were effective for learning the target grammar points. Additionally, an attitude questionnaire and interviews revealed that students held positive views towards the use of corpus-based activities.

Phoocharoensil (2012) conducted a qualitative study with Master’s students at a university in Thailand. After the students studied grammar topics such as relative clauses and conditionals using corpora, the researcher aimed to determine their attitudes towards corpus-based grammar teaching through a questionnaire and interviews. A significant majority of the students expressed satisfaction with this new teaching method, finding it useful, effective, and enjoyable. They perceived it as more productive compared to other instructional methods they had previously experienced in learning grammar.

Hanafiyeh and Keshi (2013) conducted a study to examine the effects of corpus-based instruction versus thesaurus-based instruction on the grammatical knowledge of Iranian EFL university students. The assessment of students' grammatical knowledge focused on their writing skills. In this study, the control group used an online thesaurus as a vocabulary reference tool for writing, while the experimental group utilized a web concordance. Grammatical errors were analyzed in terms of the use of nouns, adjectives, and prepositions. The results indicated that students in the concordance group demonstrated a greater reduction in grammatical errors in their writing compared to those in the thesaurus group.

Sahragard, Kushki, and Ansaripour (2013) investigated the impact of Data-Driven Learning (DDL) on the acquisition of English relative clauses by Iranian EFL learners. The study involved 40 intermediate language

learners. The researchers examined 15 writing samples from each participant and identified errors in these samples. Students in the experimental group studied relative clauses and their usage in depth using a corpus, which included translations into their first language (L1). When students struggled with inductive learning, the researcher implemented a deductive approach at periodic intervals. In contrast, students in the control group learned about relative clauses through traditional methods, relying on definitions and explanations from their teacher. During the post-test, students were required to revise their writing samples and correct mistakes related to relative clauses. The results indicated that both the traditional and the DDL methods significantly enhanced the learners' abilities to use English relative clauses effectively. Sahragard, Kushki, and Ansaripour concluded that using corpora is an effective way to address the challenges students face when learning a language.

In 2014, Smart investigated the impact of guided inductive Data-Driven Learning (DDL) on forty-nine advanced English learners at American University (Smart, 2014). The study involved three experimental groups. The first group received guided inductive, paper-based DDL instruction (DDL group). The second group underwent deductive, corpus-informed, rule-based instruction (Deductive Corpus-Informed Group - DCI), while the third group experienced traditional grammar instruction using standard teaching materials (Traditional Grammar Instruction group - TGI). The study, focused on teaching the passive voice, lasted two weeks. Instruction for the DDL group followed the 'four-Is' model proposed by Flowerdew (2009), which includes illustration, interaction, intervention, and induction. In contrast, the DCI group received a deductive approach, where corpus-informed grammar rules were presented to students who practiced these rules through the Presentation-Practice-Production (PPP) paradigm. The TGI group also employed a deductive approach, learning rules through traditional materials and practicing them via the PPP method. To measure the effectiveness of the instruction, pre-tests, post-tests, and delayed post-tests were administered to all three groups. The analysis of the results indicated that the DDL group performed the best among the three, demonstrating superior results and maintaining their achievements over time. Thus, the study concluded that inductive, paper-based DDL was highly effective for students in learning the passive voice.

In their 2015 study, Lin and Lee explored Data-Driven Learning (DDL) of grammar at a Taiwanese university (Lin and Lee, 2015). The study involved six teacher-students who were pursuing a Master's degree and had at least one year of experience teaching English. These participants underwent two training sessions, each lasting three hours, where they learned how to implement DDL and create DDL materials. After becoming familiar with the DDL method, the teachers collectively began to teach grammar to three different classes. The first group learned about the passive voice using a blended approach that combined 60% Data-Driven Learning with 40% Traditional Deductive Approach (TDA). The second group focused on relative clauses, using a different blending method that incorporated 40% DDL and 60% TDA. The third group learned lexical phrases solely through the Traditional Deductive Approach. The instruction for all groups lasted three weeks. A key distinction between the first and second groups was that the second group received step-by-step guidance and more specific questions during their DDL tasks. Post-test results showed no significant differences in the grammar acquisition outcomes among the three groups, indicating that all teaching methods were effective. The effectiveness of the treatments was evaluated from both the students' and teachers' perspectives. In the questionnaire, the groups assessed their treatments based on teaching attitudes, teaching methods, teaching content, and teaching effects. Students from two blended learning groups demonstrated an increased willingness to engage with the Data-Driven Learning (DDL) integrated treatments. The teachers were asked to keep journal entries to reflect on students' learning experiences with each treatment type. After the project, they presented evaluative reports on the treatments they conducted. The teachers indicated that DDL enhanced learners' motivation and encouraged them to become active participants in their learning. This approach created a student-centered learning environment that facilitated better learning outcomes. Additionally, interactions between students and between students and teachers were improved. However, the teachers also identified some challenges. They reported that designing DDL materials was time-consuming and added to their workload. Furthermore, extra time was necessary for students to complete DDL tasks, as many found it difficult to analyze data due to encountering unfamiliar vocabulary in the corpus. Despite these challenges, all six teachers emphasized the significance of Data-Driven Learning and regarded it as "worth the trouble" (Lin & Lee, 2015, p. 271).

Nugraha, Miftakh, and Wachyudi (2016) investigated the impact of Data-Driven Learning (DDL) on the improvement of grammatical knowledge among 30 Indonesian university students. The study was conducted over a period of two months, focusing on grammar features such as nouns, possessive pronouns, adjectives, adverbs, conjunctions, and more. The DDL approach was implemented based on the sample developed by Chujo and Oghigian (2008). The process involved several steps: First, students formed groups to observe specific grammar features using concordance printouts or a corpus concordancer, allowing them to draw their own conclusions.

Second, the teacher provided clarifications on the topics discussed by the students to guide them toward developing accurate hypotheses. Third, students practiced the targeted grammar features through homework assignments, followed by continued practice in classroom activities. To evaluate the overall DDL method and its individual components - such as the DDL worksheets, the teacher's grammar explanations, follow-up activities, and feedback from the teacher - students completed a questionnaire. The results indicated that the students had positive attitudes toward each aspect of the method, finding it both useful and enjoyable.

Elsherbini and Ali (2017) aimed to examine the effects of corpus-based activities on the grammar and vocabulary skills of EFL university freshmen, as well as their attitudes towards the use of corpora in learning. The study was conducted in Cairo, Egypt. The grammar topics covered included noun count, tenses, modals, conditionals, gerunds, infinitives, and word order. Both direct and indirect uses of a corpus were applied to the experimental group. The results revealed that the experimental group outperformed the control group in both grammar and vocabulary acquisition. Additionally, students' attitudes towards the new corpus-based instructional method were assessed through a questionnaire and interviews. Most students in the experimental group expressed positive attitudes toward using a corpus for learning grammar and vocabulary, found it to be very helpful in their language studies, and indicated a willingness to continue with corpus-based instruction in the future.

Yanto and Nugraha (2017) conducted a Data-Driven Learning (DDL) study focused on grammar with a group of twenty students learning English at an Indonesian university. The students' English proficiency levels ranged from elementary to intermediate. The target grammar points included the passive voice, conditionals, and others. The DDL procedure followed a four-step framework proposed by Chujo and Oghigian (2008). In the first stage, the grammar points were presented to the students through concordance lines. Working in groups, the students explored the patterns, made discoveries, and recorded their findings. In the second stage, each group reported its results to the class, and the teacher provided explicit explanations about the grammar features discussed. This explanation helped students verify or refine their hypotheses into more accurate ones. The third stage involved assigning follow-up exercises and production tasks as homework. Finally, in the fourth stage, during the next lecture, the teacher provided feedback on the homework, and the production practice continued. The treatment lasted for two weeks. Students' perspectives were gathered through a questionnaire and interviews. The results of the questionnaire revealed that most students found corpus-based learning to be useful. Many students also indicated that corpus-informed activities improved their grammar skills and increased their confidence in learning English grammar. However, half of the students perceived learning grammar through corpus-informed activities as challenging, finding it more difficult than learning from a conventional coursebook. More than half of the students did not express a preference for corpus-informed activities. In the interviews, students shared that inferring rules from concordance lines was difficult for them. They emphasized the importance of a teacher's guidance in understanding and analyzing the data. Additionally, some students reported that certain grammar structures are better learned through traditional textbooks rather than through corpus-based methods. Consequently, they suggested that teachers should integrate both traditional and corpus-based approaches in their teaching in a balanced manner.

Wang (2018) explored the impact of corpus-based grammar instruction on Chinese students learning English as a Foreign Language (EFL). The study involved forty freshman university students. The focus was on specific grammatical features, particularly English conjunctions such as "because," "since," and "for." The participants were divided into two groups: a control group that received traditional grammar instruction using a coursebook, and an experimental group that was taught using a corpus-based approach. The post-experimental test results indicated that the experimental group demonstrated a higher level of proficiency in understanding the target grammatical features compared to the control group. Observations of classroom activities and interviews with students revealed that they expressed positive attitudes towards learning grammar through the corpus-based method. Additionally, students reported an increase in motivation to learn grammar, achieving a deeper comprehension of the subject. The focus of learning shifted from mere memorization of forms to understanding semantic meanings and pragmatic functions. Overall, students found the corpus-based approach to be very engaging and stimulating.

Abdul-Ameer (2019) conducted a study with freshman university students in Iraq to investigate the effects of Data-Driven Learning (DDL) of English grammar on their grammatical development. The study lasted two weeks and utilized a four-step model of DDL developed by Chujo and Oghigian (2008). The analysis of the data revealed that paper-based DDL activities were effective for beginner-level English learners in acquiring grammar. This approach shifted the instructional method from a deductive to a more inductive teaching style. Additionally, the results of an attitude questionnaire indicated that students were satisfied with the integration of DDL activities in their grammar learning.

In 2019, Kim investigated the insights of teachers and students regarding corpus-based grammar instruction in a Korean EFL (English as a Foreign Language) primary school context. The study involved three teachers and eighteen students, with prepositions selected as the target grammar items. The research adopted Flowerdew's (2009) '4 Is' approach, which includes illustration, interaction, intervention, and induction. Teachers' perceptions of corpus-based grammar instruction at the elementary school level were explored through interviews, while students' attitudes were assessed using a questionnaire. The teachers reported that implicit grammar instruction often proved ineffective, leading them to emphasize the need for explicit instruction. Students performed better with this approach compared to traditional rule-based instruction. Corpus-based learning enhanced interaction in the classroom, and students enjoyed discovering patterns and drawing conclusions, even if they did not fully understand the meaning of concordance lines. However, it took time for students to become familiar with the inductive processes involved. The teachers suggested promoting hands-on use of corpus resources while avoiding excessive guidance. Students found the new methodology beneficial for learning prepositions, noting that corpus-based activities helped them retain knowledge. The discovery of rules, engagement in discussions, and idea sharing increased their motivation to learn. They perceived corpus-based grammar learning as advantageous for their reading, writing, and speaking skills. The main challenge they faced was understanding the meanings of concordance sentences, with some students expressing a desire for explanations of word meanings or translations of the concordance lines.

Data-Driven Learning is ideally suited for teaching elements of language that exist at the intersection of syntax and lexis (Bedmar, 2006). This approach addresses both lexical and grammatical collocations. Lexical collocations encompass various combinations of nouns, adjectives, verbs, and adverbs. Examples include noun + noun, noun + verb, adjective + noun, verb + noun, verb + adjective, verb + adverb, adverb + adjective, and adverb + verb (Demir, 2017). Grammatical collocations, also known as colligations, consist of combinations involving a noun, an adjective, or a verb that is followed by a preposition, an infinitive, or a clause. These can be structured as noun + preposition, noun + *to*-infinitive, noun + that-clause, preposition + noun, adjective + preposition, predicate adjective + *to*-infinitive, and adjective + that-clause (Moehkardi, 2002). Additionally, collocational verb patterns fall under the category of grammatical collocations. Moehkardi (2002) identifies nineteen types of verb patterns in the English language. Utilizing a corpus can enhance the lexico-grammatical proficiency of learners (Boontam & Phoocharoensil, 2018).

Sun and Wang (2003) conducted a study involving senior high school students in Taiwan, with a total of 81 participants. The students were divided into two groups: an inductive group and a deductive group. Initially, both groups completed a pre-test that included error correction exercises focused on English collocations. Following the pre-test, the treatments were implemented. In the inductive group, a three-step concordance-based instructional approach was used. Students explored examples of the target word, inferred rules from these instances, and corrected provided sentences based on the rules they derived. In contrast, the deductive group was given explicit rules to study, along with examples, and then corrected the provided sentences according to these rules. After the treatments, post-tests were administered, which also included error correction exercises on collocations, similar to the pre-test tasks. The results indicated that students in the experimental group performed better in learning collocations compared to those in the control group. Specifically, easier types of collocational patterns were more effectively learned through inductive, corpus-based teaching, whereas more challenging collocational patterns were equally well acquired using both the inductive and deductive approaches. Students in the study reported that easier type collocations were more effectively learned through a corpus-based approach. Thus, Sun and Wang's research demonstrated that easier collocations are more suitable for students when taught using corpus-based methods.

Koosha and Jafarpour (2006) investigated the impact of Data-Driven Learning (DDL) on the acquisition of collocations with prepositions among Iranian EFL learners. The study involved 200 university students with varying levels of English proficiency. Participants received a fifteen-session treatment focused on teaching collocations. The control group was instructed in prepositions and their collocational patterns through explicit teaching methods. In contrast, the experimental group learned collocations of prepositions using the DDL approach, which involved introducing the collocations through concordance printouts. The results from the post-test indicated that the DDL group outperformed the control group in learning collocations of prepositions.

Liu and Jiang (2009) conducted a study on corpus-based lexicogrammar learning in English as a Foreign Language (EFL) and English as a Second Language (ESL) contexts at a Chinese university and two universities in the United States. The study lasted one semester at each institution. Researchers gathered students' and teachers' perceptions of corpus-based lexicogrammar instruction through a questionnaire. The study found several positive effects of corpus-based lexicogrammar learning, including enhanced language awareness, improved understanding

of lexicogrammar patterns, a deeper critical understanding of grammar, and a more enjoyable and beneficial learning experience. Most students reported that the use of corpora was helpful for their learning, with nearly half feeling they had learned significantly from this approach. Additionally, many students expressed enthusiasm for using corpora in their future learning endeavors. Half of the students also noted that they began to see grammar and vocabulary as being more closely related than they had previously thought. However, students faced challenges in corpus-based learning, including the overwhelming amount of data, many unfamiliar words within the data, and insufficient training in how to effectively use corpora. Several strategies were suggested by both instructors and students to help manage these difficulties. One effective practice proposed by the instructors is to begin with search activities focused on deductive learning. This involves identifying instances that illustrate a given rule or pattern. By familiarizing students with the nature of the corpus in this way, their confidence and motivation improve, preparing them for inductive corpus searches. Another recommendation from the instructors is to have students conduct group corpus searches. Working in groups allows students to solve complex tasks more efficiently and quickly than they would as individuals. Additionally, group work fosters peer learning and boosts self-confidence among students. The instructors also noted that lexicogrammar items with multiple meanings are best learned through the corpus. Finally, they recommend that teachers encourage learners to use dictionaries when analyzing corpus data.

Ucar and Yükselir (2015) investigated the impact of corpus-based activities on verb-noun collocations in English as a Foreign Language (EFL) classes. The study involved 30 university students at the pre-intermediate level of English proficiency. The control group learned verb-noun collocations through traditional teaching methods using a coursebook and dictionaries, while the experimental group engaged with collocations through corpus-based materials. The results from the post-experimental test indicated that the corpus-based approach to teaching collocations was more effective than the conventional method.

Similarly, Yilmaz (2017) examined the effect of Data-Driven Learning on Turkish EFL students' acquisition of lexico-grammatical patterns in their writing. Thirty pre-intermediate English proficiency students participated in the three-week study. To address the potential challenges posed by authentic language for intermediate learners, a specially compiled corpus on various topics was used. The study focused on ten abstract nouns: pleasure, excitement, mystery, attention, permission, importance, difficulty, silence, promise, and argument. In the pre-test, students written a story incorporating the given abstract nouns. Following the pre-test, the students in the experimental group reviewed their stories for errors by consulting the corpus, whereas those in the control group used dictionaries for correction. In the post-test, students wrote new stories using the same abstract nouns. The results indicated that the experimental group demonstrated a broader range of lexico-grammatical (including collocational and colligational) patterns and exhibited greater accuracy in their linguistic usage compared to the control group.

Akinci and Yildiz (2017) explored the impact of corpus consultation on learning collocations among advanced English Language Teaching (ELT) students in Turkey, focusing specifically on verb + noun collocations. This type of collocation is categorized as a lexical collocation, and the study was considered significant alongside other Data-Driven Learning (DDL) related research. The researchers formed three instructional groups: a Data-Driven Learning group, an explicit instruction group, and a combined methods group (which utilized both DDL and explicit instruction). The treatment lasted for five weeks. The combined group experienced DDL for half the duration of the experiment, followed by explicit instruction for the second half. After the treatment, the explicit instruction group achieved the highest scores on the test, followed by the combined group, with the DDL group scoring the lowest in both the recognition accuracy of collocations and the judgment of their acceptability. Additionally, the learners' opinions about corpus consultation in learning collocations were gathered through a self-evaluation questionnaire and semi-structured interviews. The interviews revealed that students in the combined group viewed Data-Driven Learning as more effective than explicit teaching.

Girgin (2019) examined how corpus-based activities impact the acquisition of phrasal-prepositional verbs among Turkish students with upper-intermediate English proficiency. The study included fifty-eight English Language Teaching (ELT) students. A total of forty phrasal-prepositional verbs (combinations of verb + particle + preposition) were selected for instruction. The treatment consisted of six hours of instruction spread over two days, where the students learned the targeted features through corpus-based activities. The students' acquisition of phrasal-prepositional verbs was evaluated across three dimensions: form, meaning, and use. According to the post-test scores, it was found that corpus-based activities effectively taught phrasal-prepositional verbs in terms of two dimensions: form and use. The students demonstrated the ability to construct correct forms of phrasal-prepositional verbs when paraphrasing sentences. However, they struggled to grasp the metaphorical meanings of these verbs. As

a result, the study concluded that corpus-based activities were not effective in helping students learn the meanings of English phrasal-prepositional verbs.

Many scholars suggest various strategies to address the challenges associated with Data-Driven Learning (DDL). Phoocharoensil (2012) identifies two types of data presentation in DDL. The first is the presentation of raw corpus data, which may not achieve the desired outcomes due to the complexity and vast amount of information. The second approach involves a controlled exposition of data, where the teacher selects specific examples that guide students toward the correct conclusions.

Bennett (2010) proposes several strategies for adapting corpus-based activities for beginner to low-intermediate language learners. These strategies include: a) asking simple research questions, b) allowing students to find their own concordance lines, c) adapting concordance lines to match students' proficiency levels, d) including fewer lines in the materials, and e) promoting group and whole-class work.

Kim (2019), in agreement with Hunston (2002), suggests creating complete clauses or phrases from concordance lines for lower-level language learners (as illustrated in figure 7 below).

Before			After		
careful to find out all about this Ogre, went	into	the castle and asked leave to call on him. "What a	He went	into	the castle.
not think of a story to tell Frog. Then Toad went	into	the house and stood on his head. "Why are you	Then Toad went	into	the house.
HIPPOPOTAMUS AT DINNER The Hippopotamus went	into	a restaurant. He sat at his favorite table. "	The Hippopotamus went	into	a restaurant.
she buys, she puts by the savings and they go	into	the jar. When we can't get a single other coin	They go	into	the jar.
straight through the chimney. And when she went	into	the kitchen, she found her husband sitting in	When she went	into	the kitchen, she found her husband.
you I have an APPETITE!" The waiter went back	into	the kitchen. He returned carrying enough bean	The waiter went back	into	the kitchen.
is swimming," said the frog, and she jumping	into	the river. "Oh," said the young mouse, "I don't	She jumped	into	the river.
			The puppy grew	into	a dog.
			The bus went	into	a dark tunnel.

Figure 7. DDL material adaptation (Source: Kim, 2019, p. 135).

Lin and Lee (2015) propose several solutions to the challenges associated with Data-Driven Learning (DDL): a) reducing the number of concordance lines when developing DDL materials; b) shortening concordance tables; c) using more complete sentences; and d) creating specific guiding questions for students.

Sah (2015) suggests integrating Data-Driven Learning with the traditional Presentation-Practice-Produce (PPP) approach to address some of DDL's shortcomings, allowing for a more controlled learning process and better achievement of learning objectives.

This subchapter gives readers insights into various studies on the Data-Driven Learning of grammar and lexicogrammar. It focuses on study design, results, perceptions, challenges, and the authors' recommendations for overcoming them.

3. A MODEL OF DATA-DRIVEN LEARNING OF GRAMMAR USED IN THE EXPERIMENT

A model for Data-Driven Learning of grammar has been developed to apply in the study. The implementation of Data-Driven Learning of grammar was preceded by a three-week preparation stage. The preparation stage aimed to familiarize students with both indirect and direct uses of a corpus in language learning, particularly in grammar. Through the use of concordance-based materials, students were introduced to authentic input and the concepts of discovery and inductive learning. The teacher acted as a role model during the corpus-based activities.

In the preparation stage, the students in the experimental groups gained experience using the BNCweb (CQP-Edition) interface (Hoffmann & Evert, 2018) and part-of-speech (POS) or grammatical tagging. They also became acquainted with the web-based Corpus Query Processor (CQP) developed by Lancaster University (Hardie, 2012) and its tag sets, specifically UCREL CLAWS7. The Constituent Likelihood Automatic Word-tagging System (CLAWS) was created by the University Centre for Computer Corpus Research on Language (UCREL) at Lancaster University. Additionally, the learners explored certain wildcard features and engaged in basic annotation practices. Furthermore, the students in the experimental groups became familiar with the Sketch Engine corpus manager (Kilgarriff et al., 2014) and its part-of-speech tag sets. They learned to use the Corpus Query Language (CQL) tool along with various strings useful for searching complex language patterns.

The implementation stage involved studying specific grammar points by learners using the Data-Driven Learning method. In this study, several corpora were utilized by the experimental group's students.

The implementation stage involved learners studying specific grammar points using the Data-Driven Learning method. During this study, several corpora were utilized by the experimental groups' students. These included:

1. BNC*web* (CQP-Edition) (Hoffmann & Evert, 2018);
2. Corpora available on the CQP*web* server of Lancaster University (Hardie, 2012), particularly:
   - British National Corpus (XML edition);
   - Brown Family (extended);
   - BNC sampler;
   - British English 2006;
   - American English 2006;
   - Spoken English;
   - Works of Dickens.
3. Corpora on the Sketch Engine corpus manager (Kilgarriff et al., 2014):
   - British National Corpus (BNC);
   - British Academic Written English Corpus (BAWE).

During the implementation stage, several steps have been outlined throughout the process of Data-Driven Learning for a specific grammar point (Figure 8).

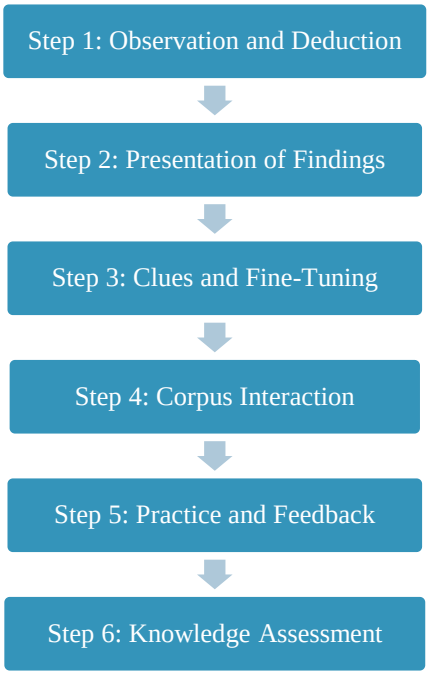

Figure 8. Steps of Data-Driven Learning of a specific grammar point.

The initial steps aim to engage students in the indirect application of corpora in the classroom. The first step, Observation and Deduction, involves introducing concordance-based materials that highlight the target grammar point. Students, whether working individually, in pairs, or in groups, examine and identify the presented grammar patterns. Through the use of guiding questions, they make assumptions, draw conclusions, and develop hypotheses.

This first step of the Data-Driven Learning process incorporates principles from the discovery learning approach and Schmidt's Noticing Hypothesis, which were discussed earlier. During this process, students employ various skills, including analytical thinking, problem-solving, teamwork, and collaboration. The guiding questions in the Data-Driven Learning (DDL) worksheet facilitate a controlled environment for learning, fitting within the framework of guided discovery.

In the second step, Presentation of Findings, students share their discoveries with the class, whether individually, in pairs, or in groups. This sharing is followed by the third step, Clues and Fine-Tuning, where the teacher provides more specific descriptions or hints to help students correct or justify their hypotheses. As a result, learners refine and adjust their assumptions and reconstruct their inferences.

The initial three steps of the Data-Driven Learning (DDL) process of grammar rely on inductive reasoning, beginning with observation and concluding with hypothesis verification. No explicit teaching strategies are employed. Instead, learners derive generalizations from specific examples, leading to the formation of assumptions that are later tested.

Each step in the DDL process promotes student-centered learning, where students actively engage in their own learning. This approach fosters an interactive learning environment. Pairing or grouping students encourages collaborative learning and helps build their confidence. Cooperative learning activities stimulate students' interest and positively influence their motivation to learn.

The subsequent step, known as Corpus Interaction, involves direct engagement with a corpus by the learners. They are encouraged to utilize corpus tools and data effectively. Under the teacher's guidance, students explore the target grammar point in authentic contexts and complete tasks that require them to apply their knowledge of the corpus.

The fourth step of the Data-Driven Learning process in grammar involves a Computer-Assisted Language Learning (CALL) technique. This step aims to create a more engaging learning environment while enhancing students' computer literacy. It fosters a student-centered atmosphere, positioning learners as researchers and teachers as facilitators. In this phase, experiential and hands-on learning practices are encouraged in the classroom, helping students to develop into self-directed learners.

Additionally, similar to the first step, authentic language input is provided to students during this phase. This exposure brings learners closer to real language, addressing their needs and interests. It facilitates a connection between learning and real-life contexts, which boosts their confidence in language awareness. Krashen's principle of Comprehensible Input (Krashen, 1985) - which suggests providing input just above learners' current level of competence - serves as a foundation for working with corpus data.

In the fifth step, known as Practice and Feedback, students engage in corpus-based tasks and activities focused on the intended grammar point. These tasks can also be assigned as homework, and learners receive feedback on their performance for each assignment.

The Data-Driven Learning (DDL) approach to grammar encompasses a series of steps that collectively reflect a constructivist method of learning. In this approach, students actively construct their own knowledge rather than passively receiving it. This construction of new knowledge is based on exploring, processing, analyzing, and interpreting corpus data, which leads to meaningful conclusions.

The final step, Knowledge Assessment, involves administering a test to evaluate students on a grammar point they have learned. A standard (non-DDL) test is used to assess learners' skills regarding a specific grammar item. This summative assessment evaluates students' performance on the test. The purpose of this step is to assess the effectiveness of the DDL method and to keep students focused on applying the corpus for targeted knowledge acquisition.

Data-Driven Learning (DDL) of grammar promotes an active learning process that encourages autonomy and helps learners develop skills for self-directed learning. The DDL approach offers a challenging yet enjoyable learning experience.

4. EXAMPLES OF ACTIVITIES IN DATA-DRIVEN LEARNING OF GRAMMAR

This section provides an overview of the steps involved in Data-Driven Learning (DDL) of grammar, along with sample activities and tasks for each step. One of the focus areas in this study is the use of comparative and superlative adjectives. The procedure for applying DDL to this grammar point is as follows:

Step 1: Observation and Deduction

Students are presented with concordance examples that illustrate the target grammar point. Some corpora allow users to download the page of query search results, or a teacher can extract a desired number of sentences from a corpus and design instructional materials accordingly. For inductive instruction of the target grammar points, teachers create DDL worksheets. These worksheets display the grammar items within a series of concordance lines. For example, Figure 9 presents an excerpt from a corpus-based worksheet designed to facilitate the inductive learning of comparative and superlative adjectives.

but they made few net gains and their 25 per cent share is **lower** than the 27 per cent they were achieving at mid-term in the last

She would not yield to his curiosity. It had been a **good** lunch and conversation had been light and easy. The Colonel had

could be credited with more than 41,667 votes each. The quota is **the smallest** number of votes to allow five candidates to be elected and is

hich a fire fighting strike is possible. Height and restricted access are **the most significant** factors of the firefighting problem. Since height in common with

sat she could see for miles the neatly sown fields, the **healthy** crops and the spotless cattle. It appeared to her as if

out the important point that movements out of and into large groups are **rarer** than movements out of and into small groups. Where nearly half the

n't I was n't looking forward to it much, it was **better** than expected. I managed to c to cover virtually the complete G

Israel. I am glad to see, however, that America is **bigger** than both of them! Wednesday night I get a phone call in

mix of Angle, Saxon, Dane and Norman: the mongrel is **healthier** than the pedigree. As such, the word English becomes abusive shorthand

and we look forward to travelling with you again. It was **the best** holiday we had had for years". Ms. A. P. (

man raving mad on the way to the churchyard. But he is **more interesting** than the others, the ones who come from the highroad to watch

of mingled tenderness and passion. I used to think: These are **the deepest** feelings I shall ever experience; this is the closest we shall

Figure 9. An excerpt from the DDL worksheet for concordance-based activities on Comparative and Superlative Adjectives (Source: BNC Sampler: *powered by CQPweb*, 2021).

In the worksheet, the concordance instances are followed by leading questions that students should focus on during their activities. These leading questions guide students toward conclusions or hypotheses regarding the specified grammar point. The leading questions are as follows:

- Can you provide examples of positive adjectives? What about comparative adjectives? Superlative adjectives?
- What do most superlatives end with?
- If an adjective ends in a single consonant, what happens to that consonant when forming the comparative or superlative form?
- If an adjective ends in *-y* and is preceded by a consonant, what happens to the *-y* when forming the comparative or superlative form of the adjective?
- What other methods of forming comparative adjectives do you notice? Which adjectives fall into this category?
- What other methods of forming superlatives do you observe? Which adjectives belong to this category?
- When is the comparative degree used?
- When is the superlative degree used?
- Do any adjectives have irregular forms in the comparative and superlative degrees? If so, group them accordingly.

Students (individually, in pairs, or groups) observe the concordance lines, address the leading questions, make inferences, and develop general premises (hypotheses) regarding the formation and use of the degrees of comparison in adjectives.

Step 2: Presentation of Findings

Students present their findings to the class.

Step 3: Clues and Fine-tuning

The teacher provides some clues or clarifications to guide learners in refining or confirming their assumptions. As a result, students draw their final conclusions.

Step 4: Corpus Interaction

Students are allowed to use a corpus to independently investigate data related to the target grammar point. The main corpora used in this study include BNC Sample and Brown Family (extended) powered by CQP web (Hardie, 2012). Students use the part-of-speech (POS) tags available online from the UCREL CLAWS7 Tagset. They receive assistance in building combinations of tags necessary for querying adjectives in degrees. Some examples of tag combinations include: (is|are) _JJR than, (is|are) the _JJT, (is|are) more _JJ than, and (is|are) the most _JJ.

Learners explore various comparative and superlative adjectives, gaining a deeper understanding of their formation, meaning, and usage. Figures 10 and 11 illustrate samples of concordance lines that display degrees of comparison of adjectives that students may obtain and analyze.

game " , one senior western diplomat commented . " But Ortega	is cleverer than	the other and he works harder . " In his quest for
classification does not necessarily mean that the quality of what is provided	is higher than	that available in an establishment with a lower classification . An explanatory
the Battle . Easy to reach by road and rail , Sedgemoor	is nearer than	you think and whatever your needs — a family holiday by the
is therefore preferable and a site sheltered from the early morning sun	is better than	one facing east . I grow the variegated form of Philadelphus coronarius
to Roznov near the Beskydy mountains . In Hungary , Lake Balaton	is larger than	Lakes Geneva or Garda and is a well developed and long established
of legend . Sometimes their life stories are so unusual that truth	is stranger than	fiction . Others are famous or notorious people who generate so many
you also were to be told but mama said nay — you	are younger than	myself , she said , and are unlikely to wed for some
treatments . Even though today 's colours , perms and styling products	are kinder than	ever before , hair is still subjected to a daily assault course
It may be noted incidentally that if in any land the poll	is higher than	the national average one or more parties in it may receive a
of this in SelecTronics ' product range . Products SelecTronics ' range	is smaller than	Franklin 's but similar in broad content . Their WordFinder products represent
erm earlier that the erm I think you said the objection site	is higher than	the Yes . site forty is higher than the road , which
said the objection site is higher than the Yes . site forty	is higher than	the road , which is which is perfectly true . Yes .
perfectly true . Yes . But in fact parts of the village	are higher than	the road , and Yes , yes of course . it 's
you 'd prefer to be alive . Life in a box	is better than	no life at all . I expect . You 'd have a
Very good very mathematical . Yes yes yeah when the first number	is smaller than	the second one . When you try and take away more than
prevent things going wrong , on the argument you know that prevention	is better than	cure and rather than trying to put problems right which have already
they heavier than air ones ? Or , or or are They	are heavier than	air and again if you follow the manufacturer 's instructions , if
on our taxpayers . Our present VAT rate , at 15% ,	is lower than	in many countries : As part of a general restructuring , it
If Sweden has lost £ 100 million because its tax rate	is higher than	ours , how much money would be attracted to Britain if our
into a domain with a partial order according to which one process	is greater than	another if it is better defined , or more predictable . If
R') . ((P' , Q' , R')	is simpler than	(P , Q , R) if each of its components
sure that any boolean appearing within the " scope " of another	is stronger than	it . The above example also illustrates the point that if ,
the workers population is working in India . I mean the poverty	is greater than	in Africa or anywhere ah in the world . you know .
You know , you only . In Kwik Save , Nescafe	is dearer than	Traidcraft coffee . Oh . Really ? Well , Traidcraft coffee 's
ca n't . So even when we buy Traidcraft , unless it	is cheaper than	it was ten years ago , it should n't really be ,

Figure 10. A sample of concordance lines for comparative adjectives (Source: BNC Sampler: *powered by CQPweb*, 2021).

war in Iraq , the Minister said : " The Iraqi people	**are the best**	to judge that . They want their country back . " Mr
for anything worth stealing . The use of " trap houses "	**is the latest**	weapon being used by police in south Leeds , as part of
on a list of 33 in Notts earmarked for closure . It	**is the nearest**	counter for residents at Brookdale Court , The Cedars and Valley Court
before you know it . I 'll tell you honestly , you	**are the best**	, nicest , grandest , sweetest girl in the world . No
research from trade union body the TUC . The increase in Wales	**is the highest**	in the UK , rising by 1.3% since last year to 12.2%
fool - you bet that Tommy sees ! " Disgrace cop PAEDOPHILES	**are the lowest**	of the low . Any police chief who disagrees should be stripped
EDWARD RUMFITT . Upton Park , E London Dear Sun GLOBAL warming	**is the greatest**	tax cow in history and abused by Labour to take more of
Armed Forces may not be the biggest , but man for man	**are the best**	. This will no longer be true , unless we change course
is a pity because of the two ways the path of publicity	**is the easiest**	to defend . The prince is known to want to contribute to
it and kids even bring them into school . The older ones	**are the worst**	, like the 17 to 19-year-olds . They 're well scary .
of the Premier League is ! David Carraway , Toxteth RAFA BENITEZ	**is the best**	manager in the world . We need to stop all these ridiculous
the conveniently pithy exegeses of Hirsi Ali and other neocon pugilists ,	**is the easiest**	and quickest way to figure it all out . Serenely , they
be " preferable " if he did n't come . But that	**is the limpest**	possible way of expressing an objection and , unsurprisingly , Mugabe dismissed
Bill . The clash of tides , known as the Race ,	**are the fiercest**	in Europe and come roaring in to explode in a frenzy on
will on the premium (60GB) model . Q. So which	**are the best**	games ? A. Super Mario Galaxy - a brilliant , all-ages adventure
A. Super Mario Galaxy - a brilliant , all-ages adventure -	**is the best**	thing to play on the Wii solo , while another recent release
) The most difficult of the terms , because historically it	**is the richest**	, is 'culture' . The critical theorists I am writing both with
encourages the growth of a fungus called sooty mould . Early summer	**is the best**	time to tackle the problem , when the newly hatched insects are
from D to Z determine the value according to colour - D	**is the clearest**	and most expensive while Z contains more traces of yellow . However
after the matches and so on . Which league do you think	**is the strongest**	in Europe at the moment ? Neil Smith , Epsom No doubt
saw this a designer handbag that 's been Botoxed ! Alex Curran	**is the latest**	famous type to fork out 1,500 for the Zagliani python-skin bag ,
The WAG was upset by the incident but unhurt . This attack	**is the latest**	in a long line of raids on Liverpool footballers ' homes .
playing fashion forward Summer in the hit TV series The OC ,	**is the latest**	in a string of Hollywood starlets to announce a new clothing collaboration
what goes on in their minds and how what you may think	**is the best**	way to help your children could be doing more harm than good
responsibility of society as a whole to make sure that being healthy	**is the easiest**	option for all of us . Healthy lifestyle education starts at home

Figure 11. A sample of concordance lines for superlative adjectives (Source: Brown Family (extended): *powered by CQPweb*, 2021)

In addition to providing a general overview of the data, specific tasks are assigned to students. Regarding the grammar point previously mentioned, one such task can be:

Use a corpus to find examples of

- Comparative adjectives formed by adding *-er*.
- Comparative adjectives formed by adding *more*.
- Superlative adjectives formed by adding *-est*.
- Superlative adjectives formed by adding *most*.
- Comparatives and superlatives derived from other roots.

Typically, the activities in the Corpus Interaction step include:

a. Corpus exploration for understanding the meaning of a specific grammatical pattern.
b. Corpus application to investigate the usage of a particular grammatical pattern.
c. Corpus exploration to compare the usages of similar or contrasting grammatical patterns.
d. Corpus application to gather a variety of instances of a given grammatical pattern.
e. Corpus exploration to resolve a specific grammatical issue.

Step 5: Practice and Feedback

Practice on the target grammar items can be carried out through both paper-based and hands-on activities. For paper-based activities, corpus-based tasks can be designed by the teacher, or relevant tasks from corpus-informed grammar textbooks that suit the students' proficiency level can be used. In Data-Driven Learning tasks, original concordance lines are preserved or adjusted as necessary.

Throughout the study, the teacher created various types of tasks using corpus data. These tasks included multiple-choice, gap-filling, matching, True/False, error correction, transformation, and both closed-ended and open-ended tasks.

Figure 12 displays an excerpt from the DDL worksheet that includes a concordance-based task focused on the comparative and superlative forms of adjectives.

as 't was really none of my business. about a month **later** than this, he told me one morning that he had had several

applicants for training than there are known opportunities. We hope the **latter** will soon catch up with the former, the advantages to the

record sales are disappointing, an internal memo disclosed. Williams ' **latest** album received mixed reviews and was overtaken in the sales charts by

ethnicity and education in East London.1 East London is one of the **last** parts of London to become gentrified and the research investigates the relationship

appropriate than a silent salute . Her husband , George Brodrick 's **elder** brother , was half blind , but refused to acknowledge the fact

as an unexpected birthday present . Eve , not quite two years **older** than me , was a shy , rather timid child , who ,

fault that was , come next morning 's debriefing? Being the **eldest** child is like being a minesweeper . But I fast-forward . I

and slyness . Seigner is the dean of the company , the **oldest** actor in point of continuous service . In that function , he

to the loch betimes to have a swim , and had walked **farther** than he had realised . He forgot to add that he had not

as Berger recently reminded us , many Palestinians today 'can go no **further** than twenty kilometres in any direction' , the casualness with which Als crosses

park of the Flying Horse and the Mercedes was parked in the **farthest** corner from the road . Even better , the driver 's side

He flung himself inside and threw his hat and gloves into the **furthest** corner , the moment the carriage began to move down the driveway

Figure 12. An excerpt from the DDL worksheet for the concordance-based task on Comparative and Superlative Adjectives (Source: British National Corpus (XML edition), 2021).

The worksheet includes a matching task that follows the concordance lines. For example,

Some adjectives have two forms of comparison. Analyze the concordance lines to match the pairs accordingly.

1.	*Later, latest* refer to	a. order
2.	*last* refers to	b. time
3.	*older, oldest* refer to	a. age and duration
4.	elder, eldest refer to	b. distinction between family members
5.	*near, nearest* refer to	a. order
6.	*next* refers to	b. distance, time

7. *farther* and *further*, *farthest* and *furthest* a. are the same in meaning

b. are different in meaning

Figure 13 illustrates an excerpt from the DDL worksheet, which features a task focused on the degrees of comparison of adjectives.

Look at the following concordance lines. They are examples of degrees of comparison of adjectives. Write *P* for the Positive, *C* for the Comparative and *S* for the Superlative.

of the running. Actually, you need to be very much better than your opponent to benefit from a superiority decision; a slight difference

pinned to her belt. The result was all she intended — glamorous, enticing, unusual. She would do her hosts credit.

Liquid, Chlorine-Free Bleach. Some of Britain 's rivers are now cleaner than they 've been for years. But a great many rivers are

essentially deterrents rather than everyday weapons. The black widow spider is the most dangerous

of all spider species and its venomous bite has been known to

' Her sister continued: In Pakistan the boys are so very obedient to their families. All right, they say, a wife

Well, here we are at last! I think I 'm more nervous than you are! Are you ready to begin? Today we 'll

vinegar. And hazelnut stuffed aubergines, mozzarella cheese, parmesan, the tiniest button mushrooms she could find. Zombie no go unless you tell

mouth. He wanted her to come out. It was a scary journey, scrabbling over that slippery great red tongue and passing through

scattered with all the motifs of the season and what could be more delightful than Wendy Phillips ' teddies, kissing under the mistletoe? Iris Bishop

no deaths. Assertions of danger are false; this treatment is safer than many prescribed drugs and unrelated to desensitisation, in which increasing doses

Figure 13. Concordance-based task on Degrees of Comparison (Source: British National Corpus (XML edition): *powered by CQPweb*, 2021).

Figure 14 shows another concordance-based task that addresses the comparative and superlative forms of adjectives.

Complete the concordance lines with the correct comparative or superlative forms of the adjectives listed below (Note: Use *than* with comparatives, *the* with superlatives)

long, near, heavy, bad, expensive, interesting, easy, little, deep, cold, attractive, important

ds are not only unhealthy, but that they are also _____ 'ordinary' foods. Nestle provi

aiters, musicians and gardeners. House slaves had _____ lives. Often they were the children of

office?" She spoke very slowly and her voice was _____ the snow outside. I looked at my wa tch

vy bags on their backs. Marina's bag was a little _____ anyone else's. 'Teddy goes everywhe

m turning more people into victims of crime. It's _____ we can do. We trust you to help us. C iti

King's thirteenth time in jail, but this jail was _____ Reidsville Prison. King was very worrie

d. It's very kind of you. I know our stay is much _____ you thought.' Diane went out to check

reported that western routes pass through some of _____ countryside around Harrogate.

wild songs, revealing at once the highest joy and _____ sadness. They would compose and si ng a

lived on the other side of town. nail's house was _____, so they decided to go there and get s

ne night, Byron suggested a game. Who could write _____ horror story? Mary's story wa

coffee fields. We helped each other. Friends were _____ money. People were good and ki

Figure 14. Concordance-based task on Comparative and Superlative Adjectives (Source: Brown Family (extended): *powered by CQPweb*, 2021).

Each task that learners complete is followed by feedback from the teacher. Feedback on classroom activities is provided immediately, while feedback on homework tasks is given individually after students submit their assignments. This feedback can take the form of qualitative comments and/or quantitative points.

Step 6: Knowledge Assessment

In this step, a test is administered to evaluate the knowledge students have acquired regarding the target grammar point covered in the earlier stages of the Data-Driven Learning process. This assessment may involve non-DDL ready-made or teacher-developed tests. The teacher grades the test, and the correct answers are totaled.

Thus, this subchapter reviewed the Data-Driven Learning process of grammar, using one target grammar point from the study as an example.

5. STUDY 1 (EXPERIMENTAL STUDY): MEASURING THE DEVELOPMENT OF UNDERGRADUATE STUDENTS' GRAMMAR SKILLS THROUGH DATA-DRIVEN LEARNING OF GRAMMAR

A. Objectives, Settings and Participants

The objectives of the study were to:

1. Evaluate the effectiveness of Data-Driven Learning of Grammar;
2. Evaluate the effectiveness of a mixed-methods approach to grammar teaching (Data-Driven Learning combined with the traditional (deductive) teaching approach);
3. Compare the effectiveness of Data-Driven Learning of Grammar with the traditional (deductive) approach to teaching grammar.
4. Compare the effectiveness of Data-Driven Learning with the mixed-methods approach to teaching grammar (Data-Driven Learning combined with the traditional (deductive) teaching approach);
5. Investigate the effectiveness of Data-Driven Learning in terms of knowledge retention.

The study was conducted at a state university in Tbilisi, the capital of Georgia. It took place during the second semester (spring) of the 2020-2021 academic year and the first semester (fall) of the 2021-2022 academic year.

Georgia adopts the European Credit Transfer and Accumulation System (ECTS) for its academic programs. All first-cycle degree programs in the country consist of 240 ECTS credits, while second-cycle degree programs comprise 120 ECTS credits, and third-cycle programs include 180 ECTS credits. The curriculum for four-year first-cycle degree programs includes the teaching of foreign languages during the first and second years of study. This study focused on first-year undergraduate students at a particular university who are majoring in English. A quasi-experiment differs from a randomized or true experiment; in a randomized experiment, participants are randomly assigned to experimental (treatment and control) groups, whereas a quasi-experiment uses pre-existing groups (Williams, 2007). The research sample was selected using a convenience sampling method, meaning the sample was drawn based on the researcher's accessible subjects. Additionally, purposive sampling was employed to target the population based on the study's objectives. In this case, the aim was to ensure homogeneity among students in terms of their English language proficiency. The final sample comprised three groups from the Faculty of Education Sciences at this university. The groups were homogeneous in terms of English proficiency, all at the B1 level. The English proficiency levels of the students were determined by the faculty administration based on their results in the Unified National Exams, which are common national university entrance examinations established by the Ministry of Education, Science, Culture, and Sports of Georgia. Participants were randomly assigned to one of three groups: a control group, experimental group 1, and experimental group 2. In total, there were 71 participants, consisting of 66 females and 5 males. Educational programs in the field of Education Sciences tend to have more female than male students, and the university where the experimental study was conducted is no exception. Consequently, female students outnumbered male students, although the ratios of males to females were consistent across all three groups. The ages of the participants ranged from 18 to 20 years.

Table 1. Distribution of students in groups based on number and gender.

Groups	N	Gender	
		Female	Male
Experimental group 1	24	22	2
Experimental group 2	25	23	2
Control group	22	21	1

B. Ethical Issues

At the outset, permission was obtained from the university administration to conduct the experiment. Before starting, participants were provided with a general overview of the teaching approach and the nature of the experiment. Through the informed consent form, participants were informed that they would not be subjected to any harm and that they could withdraw from the study at any stage. They were assured that their anonymity and confidentiality would be maintained. Participants were also informed that if the data collected from the experiment

and survey were to be reported or published, assigned codes would be used to protect their identities. Additionally, they were made aware that the information gathered would be used solely for this research and would not be disclosed to any third party. Signed consent forms were obtained from all participants.

As previously mentioned, three groups were selected to participate in the experiment. Each group had two contact hours (one contact hour = 50 minutes) of English instruction twice a week. Out of these two hours, 35 to 40 minutes were designated for teaching grammar as part of the course. Therefore, all three groups devoted the same amount of time to studying grammar during the lectures. The grammar topics covered were identical for all groups, but the teaching methods varied for each group. The teaching procedures for each group are described in the subsections below.

C. Teaching the Control Group

In Georgia, grammar is traditionally taught using the deductive teaching method, which involves explicitly teaching grammatical rules first, followed by practice and application. This approach is rule-driven, progressing from general principles to specific examples. English teaching textbooks generally include the conventional types of grammar exercises: controlled exercises (Fill-in-the-blank, Multiple-choice questions, Sentence transformation, Matching exercises), semi-controlled exercises (Sentence completion, Error correction, Sentence reordering) and free exercises (guided writing).

The control group consisted of students receiving traditional instruction in grammar. During the first half of the experiment (spring semester of the 2020-2021 academic year), these students were taught using *Laser B1* (Mann & Taylore-Knowles, 2016a). In the second half of the experiment (fall semester of the 2021-2022 academic year), they used *Laser B1+* (Mann & Taylore-Knowles, 2016b). The specific grammar points covered in each textbook were as follows:

- Articles;
- Countable and uncountable nouns;
- Comparatives and superlatives of adjectives;
- Tenses (all); *Will & Be going to*; *Would/used to/be used to*; stative verbs;
- Infinitives and -*ing* forms;
- Modals; modal perfect;
- Relative clauses;
- Result clauses (*so, such, too, enough*);
- The passive;
- Conditionals; Expressions *unless, in case, as long as*
- Reported speech; Reported questions;
- Question tags;
- The Causative,
- Direct and indirect objects;
- Conjunctions (*although, even though, despite, in spite of*);
- Expressions *Wish* and *if only*;
- Expressions *Prefer, would rather, had better.*

The grammar points in *Laser B1* and Laser B1+ are the same; however, the level of difficulty in *Laser B1+* is slightly higher than in *Laser B1*. The control group received explicit instruction from the teacher, who employed a deductive approach to teach the grammar points outlined in the textbook. The teacher provided explanations and rules regarding the formation and usage of target grammar points, and the students practiced applying these rules through various examples. Initial practice occurred during the lecture following the teacher's explanations, with additional exercises assigned as homework.

The exercises included in the textbooks consisted of activities such as filling in the correct forms, completing sentences, choosing the appropriate options, matching to create sentences, correcting sentences, transforming sentences, and other traditional grammar exercises. During the lecture, the teacher provided oral feedback to the students, while written feedback in a mixed format (including both comments and scores) was given for each homework assignment.

D. Teaching the Experimental Group 1

Experimental group 1 (the DDL group) received grammar instruction based on the Data-Driven Learning (DDL) model specifically developed for this study. This group engaged in entirely inductive teaching. The grammar points covered in experimental group 1 were the same as those taught to the control group.

Students in experimental group 1 went through six steps of the Data-Driven Grammar Learning model. In the first step, Observation and Deduction, students worked on concordance-based tasks related to the target grammar point, which had been designed by the teacher in advance. Through guided questions, they developed hypotheses about the formation and usage of the grammar point. During the second step, Presentation of Findings, students shared their findings and inferences with the class. The third step, Clues and Fine-Tuning, focused on refining or confirming these hypotheses. In the fourth step, Corpus Interaction, students consulted a corpus and completed small assignments assigned by the teacher. The fifth step, Practice and Feedback, was entirely dedicated to practicing the target grammar point through DDL tasks created by the teacher.

The teacher developed various types of Data-Driven Learning (DDL) tasks for classroom and homework activities. These tasks included multiple-choice questions, matching exercises, true/false questions, gap-filling activities, error correction, transformations, and close or open-ended tasks. Generally, the tasks were created based on original concordance lines to ensure authenticity. Occasionally, these lines were slightly revised to better suit the students' language proficiency levels. After completing each DDL task, students received feedback from the teacher. Oral feedback was provided during classroom activities, while written feedback - consisting of comments and points - was given for homework assignments. Finally, in the Knowledge Assessment step, a test was administered to evaluate students' knowledge gained through the Data-Driven Learning process.

E. Teaching the Experimental Group 2

Experimental group 2 (DDL + traditional teaching) utilized a mixed-methods approach to teaching grammar. This group combined the Data-Driven Learning (DDL) method with a deductive approach to teaching EFL (English as a Foreign Language) grammar. The target grammar points for experimental group 2 were the same as those for the control group and experimental group 1.

The teaching process began with the Observation and Deduction step (step 1), where students explored the provided concordance lines and worked on leading questions. This was followed by the Presentation of Findings step (step 2), in which students shared their results with the class. Next was the Clues and Fine-tuning step (step 3), where students refined the hypotheses they had formulated in the previous steps. In the Corpus Interaction step (step 4), students engaged with corpora and addressed specific questions. These four steps make up the inductive learning procedures. After completing these steps, the instruction for experimental group 2 shifted to a deductive approach, incorporating explicit instruction into the process. The teacher then provided all the relevant rules and explanations regarding the target grammar point. As a result, students in experimental group 2 experienced both Data-Driven Learning of grammar and explicit instruction familiar to the control group.

The final step for Experimental Group 2 was the Practice and Feedback step (Step 5) of the model. In this step, both Data-Driven Learning (DDL) tasks and traditional exercises were utilized. The exercises aligned with the target grammar points were taken from two corpus-informed books: *Grammar and Beyond 3* (Blass, Iannuzzi, Savage, Gordon, & Reppen, 2013) and *Touchstone 4* (McCarthy, McCarten, & Sandiford, 2014). These books are designed for students at the B1-B2 levels of the Common European Framework of Reference (CEFR). Both *Grammar and Beyond* and *Touchstone* series include real-world data from the Cambridge International Corpus. Each grammar point is introduced with explanations, including rules, examples, and tables, followed by traditional exercises that utilize authentic language examples. The ratio of teacher-developed DDL tasks to corpus-based traditional exercises was maintained at fifty-fifty (50:50). Feedback, both qualitative and quantitative, was provided for all DDL tasks and exercises taken from the books.

F. Data-Collection Tools and Procedures

Tests were used as research tools to achieve the study's objectives. The researcher developed five tests: a pre-test, two midterm assessments (while-test 1 and while-test 2), a post-test, and a delayed test. These tests were designed to evaluate the progress of the students participating in the experiment. Each test consisted of 50 items, which included various question types: 20 multiple-choice questions, 10 matching questions, 10 questions requiring the selection of

the best options, and the final 10 questions aimed at arranging words in the correct order to form sentences. Each correct answer was valued at 2 points.

To ensure the validity of the tests, the researcher assessed their content validity, face validity, and construct validity. Three qualified experts in the field of English Philology evaluated these three types of validity for each test. Based on the experts' recommendations, some items on the tests were refined.

To verify the reliability of the research tools, the tests were piloted with a group of 10 students. We measured the reliability of the tests using Cronbach's Alpha analysis. For each test and the group of ten students, we calculated the Pearson correlation and the significance (Sig.) (2-tailed) values. Table 2 presents the Pearson correlation and significance values for the tests.

Table 2. Reliability of the experimental research tools

Test	Pearson correlation	Sig. (2-tailed)
Pre-test	0.994	0.000
While-test 1	0.990	0.000
While-test 2	0.993	0.000
Post-test	0.988	0.000
Delayed test	0.987	0.000

The results of the five tests indicated strong correlations, with each Pearson correlation exceeding 0.8 and significance levels below 0.01. This suggests that the test-retest results were highly reliable and that the overall findings are statistically significant. Additionally, the pilot results showed that there were no questions answered correctly by 90% or more of the students, nor were there any questions that 90% or more of the students failed. Therefore, there was no need to replace any of the items in the tests.

The experiment was conducted during the second semester of the 2020-2021 academic year and the first semester of the 2021-2022 academic year. Students in the Control Group, as well as Experimental Group 1 and Experimental Group 2, took a pre-test during the first week of March 2021. They subsequently completed a mid-term test (while-test 1) in the first week of June 2021. A second mid-term test (while-test 2) was administered in the first week of November 2021, followed by a post-test in the last week of December 2021. Finally, the students took a delayed test in the last week of January 2022. The tests were graded by the researcher, and the quantitative data collected were analyzed using the SPSS version 22.0 statistics program.

G. Experiment Results and Discussions

To evaluate the improvement in grammar skills among the control group, experimental group 1, and experimental group 2, quantitative data from several tests - namely the pre-test, while-test 1, while-test 2, post-test, and delayed test - were collected and analyzed throughout the study. The results for the control group (CG) are presented in Table 3, for experimental group 1 (EG1) in Table 4, and for experimental group 2 (EG2) in Table 5. Each test was scored out of 100 points.

To ensure that there were no significant differences in grammar skills among the groups prior to the experiment, a pre-test was administered. The results indicated that the average scores (mean pre-test scores) were similar across all three groups: the control group had a mean score of 46.23, experimental group 1 had a mean of 45.71, and experimental group 2 had a mean of 45.

The pretest results for the control group indicated that the median score was 46, the mode (the most frequently occurring score) was 36, and the standard deviation (a measure of the spread of scores) was 9.28. For the mean to be considered reliable, it should be close to both the median and the mode. Additionally, the standard deviation should be less than 10% of the measurement scale, which is around 10 for a 100-point assessment system. In this case, the control group's pretest mean score is reasonably close to both the median and the mode, and the standard deviation is below 10. These results support the reliability of the mean score.

The skewness value, which measures the symmetry of the distribution, was -0.01, while the kurtosis value, which assesses the presence of outliers, was -1.06. Both skewness and kurtosis fall within the range of -3 to 3, indicating a normal distribution. This confirms that the results are reasonably reliable.

Table 3. CG progress tests mean results.

Student	Pre-test	While-test 1	While-test 2	Post-test	Delayed test
1	38	44	48	59	53
2	53	55	62	76	75
3	32	36	41	55	52
4	58	61	67	64	63
5	52	53	65	75	71
6	45	47	55	66	62
7	47	49	55	69	64
8	44	46	51	68	61
9	50	58	57	66	63
10	51	53	60	68	62
11	54	56	60	72	64
12	52	54	61	70	67
13	35	37	41	53	52
14	61	63	70	84	82
15	55	59	66	82	83
16	36	41	46	43	42
17	36	38	45	53	51
18	62	63	69	81	80
19	31	33	40	58	56
20	40	42	50	62	61
21	44	47	56	69	64
22	41	43	51	63	61
Mean	46.23	49	55.27	66.18	63.14
Median	46	48	55.5	67	62.5
Mode	36	47	41	53	61
Stand. Dev.	9.28	9.10	9.33	10.22	10.39
Skew-ness	-0.01	-0.06	-0.109	-0.229	0.286
Kurto-sis	-1.06	-1.11	-1.073	-0.011	0.069

Table 4. EG1 progress tests mean results.

Student	Pre-test	While-test 1	While-test 2	Post-test	Delayed test
1	35	46	55	62	64
2	39	50	57	70	68
3	45	52	66	73	72
4	51	59	71	81	80
5	36	46	55	61	58
6	34	42	51	64	65
7	29	37	47	67	66
8	53	56	67	81	79
9	44	51	59	75	73
10	53	62	78	91	89
11	50	59	71	88	87
12	59	67	71	85	87
13	42	46	59	81	80
14	45	50	62	80	78
15	43	53	67	70	69
16	49	55	70	91	89
17	55	61	73	89	88
18	52	58	70	92	91
19	39	47	54	66	64
20	37	48	59	73	72
21	66	71	82	79	81
22	57	64	79	95	95
23	51	55	60	92	88
24	33	43	42	66	63
Mean	45.71	53.25	63.54	78	76.92
Median	45	52.5	64	79.5	78.5
Mode	39	46	59	81	64

Table 5. EG2 progress tests mean results.

Student	Pre-test	While-test 1	While-test 2	Post-test	Delayed test
1	44	48	57	56	55
2	34	49	47	59	56
3	35	46	58	76	75
4	33	43	49	65	64
5	46	53	67	77	73
6	61	73	82	99	94
7	43	59	60	83	81
8	30	38	49	71	67
9	63	69	83	82	79
10	34	46	52	78	75
11	45	54	72	93	89
12	40	47	64	81	78
13	60	69	80	99	97
14	38	51	61	85	81
15	51	58	75	93	94
16	30	39	50	74	72
17	40	52	60	74	71
18	36	48	59	64	61
19	51	54	66	85	84
20	53	60	71	95	93
21	50	55	64	82	80
22	51	56	72	94	92
23	54	64	75	97	94
24	51	54	74	96	92
25	52	64	77	75	77
Mean	45	53.96	64.96	81.32	78.96
Median	45	54	64	82	79
Mode	51	46	49	74	94

Stand. Dev.	9.35	8.36	10.24	10.75	10.75
Skewness	0.149	0.218	-0.154	-0.11	-0.043
Kurtosis	-0.563	-0.331	-0.477	-1.33	-1.292

Stand. Dev.	9.77	9.14	10.89	12.51	12.33
Skewness	0.12	0.302	-0.04	-0.309	-0.325
Kurtosis	-1	-0.312	-1.07	-0.703	-0.739

The pre-test results for EG1 are as follows: the mean is 45.71, the median is 45, the mode is 39, the standard deviation is 9.35, the skewness is 0.149, and the kurtosis is -0.563. For EG2, the pre-test results are: the mean is 45, the median is also 45, the mode is 51, the standard deviation is 9.77, the skewness is 0.12, and the kurtosis is -1.

Overall, the mean scores of all three groups are close to their respective medians, with standard deviations around 10. Additionally, the skewness and kurtosis values range between -3 and 3. These findings indicate the reliability of the results.

Figure 9 illustrates the mean scores from all five tests for each group.

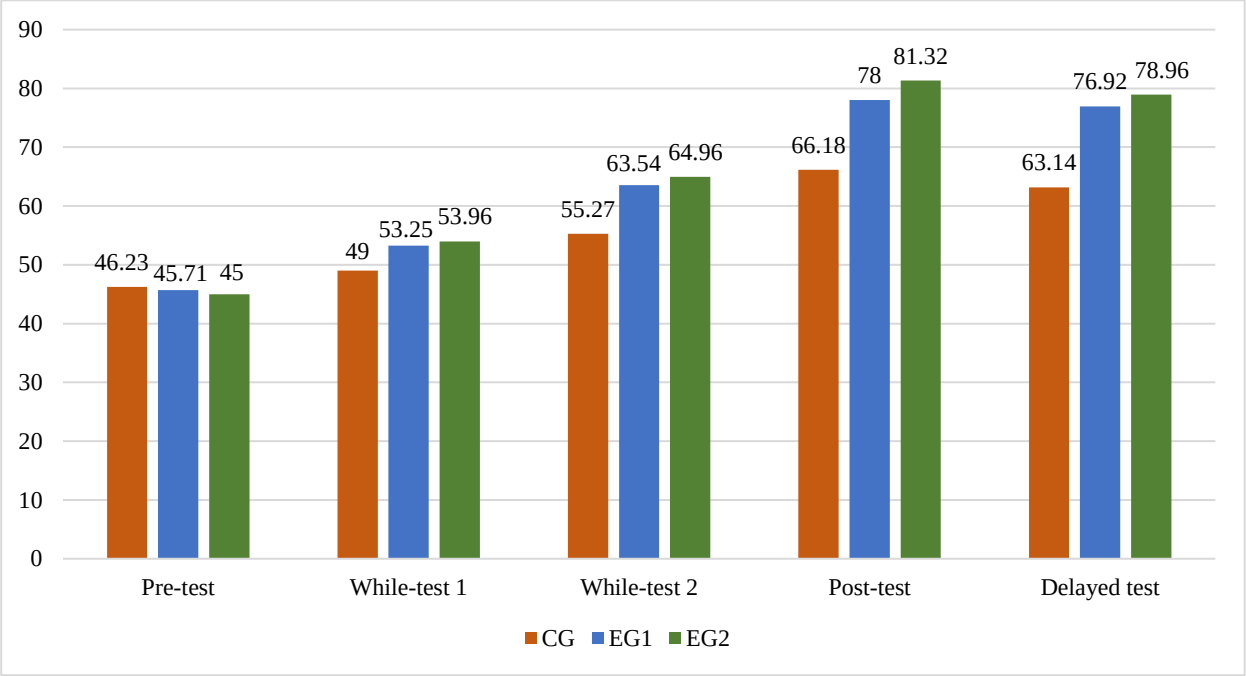

Figure 15. Mean scores of progress tests of CG, EG1 and EG2.

The results of the study showed that the average scores increased in all three groups. For the control group (CG), the mean scores were as follows: the pre-test score was 46.23, while-test 1 was 49, while-test 2 was 55.27, and the post-test score was 66.18. For experimental group 1 (EG1), the mean scores were: pre-test at 45.71, while-test 1 at 53.25, while-test 2 at 63.54, and post-test at 78. For experimental group 2 (EG2), the mean scores were: pre-test at 45, while-test 1 at 53.96, while-test 2 at 64.96, and post-test at 81.32.

To analyze the results in more detail, we observed the gains from the pre-test to test 1. The gain for the control group was 2.77, for EG1 it was 7.54, and for EG2 it was 8.96. Thus, experimental group 2 showed the greatest improvement, followed by experimental group 1, which also outperformed the control group.

Between while-test 1 and while-test 2, the gain for the control group (CG) was 6.27, while experimental group 1 (EG1) had a gain of 10.29, and experimental group 2 (EG2) had a gain of 11. The results from the second mid-test

indicated that experimental group 1 achieved significant success compared to the control group, although it fell slightly short of the performance of experimental group 2.

Soon after the conclusion of the experiment, the post-test results were analyzed. After the while-test 2, the control group (CG) improved their results by 10.91, the first experimental group (EG1) improved by 14.46, and the second experimental group (EG2) improved by 16.36. The progress made by experimental group 2 is particularly notable, while experimental group 1 also achieved significant success. The control group also demonstrated some improvement during the course of the experiment.

A month after the post-test, a delayed test was conducted to determine which group retained knowledge more effectively. In the delayed test, the control group (CG) saw a mean decrease of 3.04, the first experimental group (EG1) experienced a mean decrease of 1.08, and the second experimental group (EG2) showed a mean decrease of 2.36. Therefore, experimental group 1 (the pure DDL group) demonstrated better knowledge retention compared to experimental group 2 (the mixed-methods group).

Overall, from the pre-test to the delayed test, the final increase for the control group was 16.91. In comparison, experimental group 1 experienced a final increase of 31.21, while experimental group 2 had a final increase of 33.96. This means that, in terms of overall growth, experimental group 1 surpassed the control group by 14.3 mean points, and experimental group 2 outperformed the control group by 17.05 mean points. Additionally, experimental group 2 exceeded experimental group 1 by 2.75 mean points.

To determine whether the differences in results among the groups were statistically significant, paired-samples *t*-tests were conducted. Table 6 summarizes the *t*-test analysis results for the control group (CG) and experimental group 1 (EG1).

Table 6. Summary of the *t*-test results comparing CG and EG1.

Test	Control	Experimental 1
Pre-test	46.23	45.71
While-test 1	49	53.25
While-test 2	55.27	63.54
Post-test	66.18	78
Delayed test	63.14	76.92

The table below shows the results of the paired samples test for the control group and the experimental group 1.

Table 7. Paired samples test for CG and EG1.

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1 Control group - Experimental group 1	5.8224E1	11.63214	3.67841	49.90287	66.54513	15.829	9	.000

For the first pair (Control Group [CG] and Experimental Group 1 [EG1]), a 95% confidence interval for the difference was calculated. The results indicated a significant difference in the scores between the control group

(M=59.7, SD=11.8) and experimental group 1 (M=1.5, SD=0.52), with t=15.8, df=9, Sig. (2-tailed)=0.000. Since the significance (2-tailed) value, also known as the p-value, is less than 0.05, this indicates that the results are statistically significant. Therefore, the difference in achievements between the control group and experimental group 1 in this study is statistically meaningful.

Table 8 displays the results of the *t*-test analysis for the control group and experimental group 2. The following table presents the results from the paired samples test for the control group and experimental group 2.

Table 8. Summary of the *t*-test results comparing CG and EG2.

Test	Control Group	Experimental Group 2
Pre-test	46.23	45
While-test 1	49	53.96
While-test 2	55.27	64.96
Post-test	66.18	81.32
Delayed test	63.14	78.96

Table 9. Paired samples test for CG and EG2.

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 2 Control group – Experimental group 2	5.8902E1	12.63878	3.99673	49.86076	67.94324	14.738	9	.000

For the second pair of groups (CG & EG2), a 95% confidence interval was calculated for the difference in scores. The results indicated a significant difference between the control group (M=60.4, SD=12.82) and experimental group 2 (M=1.5, SD=0.53). The t-value was 14.73, with 9 degrees of freedom, and the significance level was p = 0.000, which is less than 0.05. This indicates that the difference in achievements between the control group and experimental group 2 is statistically significant.

Table 10 presents the *t*-test analysis results for experimental group 1 and experimental group 2.

Table 10. Summary of the *t*-test results comparing EG1 and EG2.

Test	Experimental Group1	Experimental Group2
Pre-test	45.71	45
While-test 1	53.25	53.96
While-test 2	63.54	64.96
Post-test	78	81.32
Delayed test	76.92	78.96

The table below shows the results of the paired samples test for experimental group 1 and experimental group 2.

Table 11. Paired samples test for EG1 and EG2.

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 3 Experimental group 1 – Experimental group 2	6.2662E1	14.12323	4.46616	52.55885	72.76515	14.030	9	.000

In the analysis of the third pair (EG1 and EG2), a 95% confidence interval was established for the difference. The results indicated a statistically significant difference between EG1 and EG2, with values of t=14.03, degrees of freedom (df)=9, and a significance level (2-tailed) of p=0.000.

In summary, the results from the paired samples tests demonstrated significant differences between the mean scores of the control group (CG) and both EG1 and EG2, as well as between EG1 and EG2.

The results indicate that both experimental groups outperformed the control group. Specifically, the achievement difference between experimental group 1 and the control group was 14.3 mean points, while the difference between experimental group 2 and the control group was 17.05 mean points. Additionally, the achievement difference between experimental group 2 and experimental group 1 was 2.75 mean points. Overall, both experimental groups performed better than the control group, with experimental group 2 showing the greatest improvement.

The results of this experimental study align with previous studies indicating that Data-Driven Learning (DDL) has a significant advantage over the traditional teaching method in improving the grammar skills of EFL university students.

Specifically, the findings support Rapti's (2010) study, which demonstrated that Data-Driven grammar teaching - with the use of concordance-based materials - had a more pronounced impact on students' final performance in grammar tests than conventional deductive teaching methods. Rapti's study lasted five months and focused on various grammar items, including tenses in both active and passive voices, conditionals, gerunds, infinitives, causatives, direct and indirect speech, and modals.

Additionally, this experimental study shares similarities with Smart's (2014) research conducted in an EFL university context. Smart's study revealed that among three treatment groups, the guided inductive DDL group achieved the highest scores compared to a deductive corpus-informed group and a traditional grammar instruction group. Although Smart's study was conducted over a shorter duration of three weeks, the DDL method proved effective even in that limited timeframe.

The results of the current experimental study align with the findings of Elsherbini and Ali (2017), who discovered that using corpus-based activities was more effective for developing grammar skills among tertiary education EFL (English as a Foreign Language) students compared to traditional instruction. Their research focused on teaching various grammar topics, including singular and plural nouns, tenses, infinitives, gerunds, conditionals, modal verbs, and word order. The effectiveness of corpus-based activities in grammar instruction was also evident in their one-semester study.

Additionally, the current study's results support the conclusions drawn by Wang (2018). Wang's research indicated that a corpus-based approach to teaching grammar had a more positive impact on the development of grammar skills

among EFL university students than conventional explicit teaching methods. This study also took place over a semester and highlighted the advantages of corpus-based grammar instruction.

A similar study was conducted by Girgin (2011), which examined the impact of corpus-based activities on the grammar skills development of EFL university students. The findings indicated that these corpus-based activities had an effect comparable to that of traditional coursebook grammar exercises. Specifically, students in the experimental group achieved similarly high results as those in the control group when learning five target grammar structures. Furthermore, knowledge retention was comparable between both groups. While Girgin's study demonstrated the effectiveness of the Data-Driven Learning (DDL) approach, it did not show any superiority over explicit instruction. It is important to note that the duration of Girgin's study was six weeks, and it is possible that longer-term application of corpus-based activities in grammar teaching could yield more significant results.

The current experimental study shares some similarities with, but also presents inconsistencies to, the findings of Lin and Lee (2015). In their research, Lin and Lee investigated three experimental groups: 1) a Data-Driven Learning (DDL) group with inductive teaching, 2) a blended method group that combined 60% DDL with 40% traditional deductive approaches, and 3) another blended method group that integrated 40% DDL with 60% traditional deductive approaches. Their results indicated no significant differences in the achievements of these three groups. Additionally, the brevity of their study, lasting only three weeks, sets it apart from the current research.

In contrast, the experimental study conducted by Smart (2014) included three groups: 1) a DDL group with inductive teaching, 2) a Deductive, Corpus-Informed (DCI) group, which utilized corpus-based materials in a deductive manner, and 3) a traditional teaching group. The findings revealed that the DDL group achieved the highest scores among the three groups. This outcome demonstrates the effectiveness of the Data-Driven Grammar Learning method, aligning with the results of the present study, even though Smart's research also had a short duration of three weeks.

H. Conclusion to Study 1

The results of the experimental study indicated that the control group students' scores improved from an average of 46.23 in the pre-test to 66.18 in the immediate post-test, but then decreased to 63.14 in the delayed test. In contrast, the scores of Experimental Group 1 (the DDL group) increased from an average of 45.71 in the pre-test to 78 in the immediate post-test, with a slight decrease to 76.92 in the delayed test. Experimental Group 2 (the mixed-methods group) showed an improvement from an average of 45 in the pre-test to 81.32 in the immediate post-test, followed by a decrease to 78.96 in the delayed test. The final score gains were as follows: the control group had a gain of 16.91, Experimental Group 1 had a gain of 31.21, and Experimental Group 2 had a gain of 33.96. This shows that the scores of both experimental groups improved significantly more than those of the control group. These results highlight the greater effectiveness of the Data-Driven Grammar Learning method and the mixed-methods grammar teaching approach (DDL combined with traditional teaching) compared to the traditional deductive approach to grammar instruction.

To compare the two experimental groups, experimental group 2 showed an overall growth advantage of 2.75 over experimental group 1. This suggests that the mixed-methods approach was slightly more effective than the pure Data-Driven Learning (DDL) method, although the difference was minimal. However, it is important to note that the decrease in performance from the post-test to the delayed test was greater for experimental group 2, indicating that experimental group 1 demonstrated a better ability to retain knowledge. In conclusion, the study indicated that both the Data-Driven Grammar Learning method and the mixed-methods approach (DDL combined with traditional teaching) were more effective in improving grammar skills among English as a Foreign Language students compared to traditional teaching alone. The mixed-methods group (DDL plus traditional teaching) achieved slightly better results than the pure DDL group; however, the pure DDL group displayed superior knowledge retention compared to the mixed-methods group. Overall, both forms of Data-Driven Learning (either in isolation or combined with traditional instruction) had a more favorable impact on undergraduate EFL students' grammar skills than the traditional approach.

6. STUDY 2 (SURVEY RESEARCH): EXPLORING EXPERIMENTAL GROUPS STUDENTS' ATTITUDES TOWARDS DATA-DRIVEN LEARNING OF GRAMMAR

A. Objectives, Participants, Data-Collection Tool and Procedures

The objectives of the study were to:

1. Explore the attitudes of students in the experimental groups toward Data-Driven Learning of Grammar, focusing on various aspects such as motivation, involvement, discovery learning, classroom interaction, authentic input, skills-oriented learning, autonomous learning, cooperative and collaborative learning, computer-assisted grammar learning, organization and clarification of DDL tasks, hypothesis formation, feedback provision, knowledge retention, corpus interaction, data-analysis, etc.
2. Determine the preferences of the experimental groups' students regarding the method of teaching English as a Foreign Language (EFL) grammar.
3. Identify the reasons behind the experimental groups' students' preference for the traditional method of grammar teaching, if applicable.

A survey was used as the research method for this study. Two separate surveys were conducted to meet the study's objectives. The participants included students from both experimental groups: the DDL group and the mixed-methods group (which combined DDL with traditional teaching). In total, 49 students participated in the surveys.

The questionnaire designed to achieve the first two objectives of the study consisted of 28 items (see Appendix 1). The questions were developed based on an analysis of the literature conducted within the scope of this monograph. Out of the 28 questions, 21 addressed students' perceptions of various aspects and features of the Data-Driven Learning method. Six questions focused on the potential difficulties students might encounter regarding different components of the method, and the final question asked about the students' preferred approach to learning grammar.

The questionnaire consisted of twenty-seven multiple-choice questions based on a five-point Likert scale. Twenty-one items required respondents to indicate their level of agreement, ranging from 1 (Strongly disagree) to 5 (Strongly agree). Six items assessed the level of difficulty, with responses ranging from 1 (Very difficult) to 5 (Very easy). The final item was a multiple-choice question.

To ensure the validity of the questionnaire, its content, face, and construct validity were evaluated by three experts in the fields of English Philology and Education Sciences. To assess the reliability of the questionnaire, a pilot study was conducted with a group of 10 students. Prior to this, the Data-Driven Learning (DDL) method was implemented for two weeks in one of the researcher's General English classes to familiarize the students with the method. After this period, the questionnaire was administered to the 10 students to evaluate its reliability. Reliability was measured using Cronbach's Alpha analysis, which involved calculating the Pearson correlation and the significance (2-tailed) values. The results of the pilot study showed that in all ten cases, the Pearson correlation was greater than 0.8, and the significance level was below 0.01. These findings indicate a strong correlation between the test-retest results, the results were statistically significant, and the questionnaire demonstrated reliability.

The survey was conducted using Google Forms during the same week that the students in the experimental groups completed the delayed test. The questionnaire aimed to achieve the third objective of the study and included a primary question about students' preferences for the grammar teaching method (see Appendix 2). The second question was a follow-up for those students who preferred the traditional approach to grammar learning. This open-ended question asked them to provide reasons for their preference for the traditional method over the Data-Driven Grammar Learning method. The survey was administered one week after the delayed test was taken by the experimental groups.

B. Results and Discussions

Table 12 presents the frequency and descriptive analysis of students' attitudes toward Data-Driven Learning in EFL grammar across the experimental groups. Using the SPSS v.22.0 statistics program, the mean, median, mode, standard deviation, skewness, and kurtosis were calculated for the responses to the first questionnaire.

Table 12. Frequency and descriptive analysis of students' perceptions of Data-Driven Learning of grammar.

Frequency and Descriptive Tables		Frequency					Descriptive Statistics					
Item	Statement	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree	Mean	Median	Mode	Standard deviation	Skewness	Kurtosis
1	The Data-Driven Grammar Learning method is an interesting and motivational approach to learning grammar.	0 % 0	2 % 1	4 % 2	35 % 17	59 % 29	4.51	5	5	0.68	-1.48	2.62
2	I was more engaged in my grammar study through the DDL method than the traditional (deductive) teaching method.	0 % 0	0 % 0	6 % 3	49 % 24	45 % 22	4.39	4	4	0.6	-0.43	-0.61
3	Learning grammar through the DDL method was enjoyable.	0 % 0	2 % 1	6 % 3	40 % 19	53 % 26	4.43	5	5	0.71	-1.21	1.59
4	Discovering grammatical features on my own was a more valuable experience than simply receiving ready-made knowledge from my teacher.	4 % 2	6 % 3	10 % 5	43 % 21	37 % 18	4.02	4	4	1.05	-1.28	1.42
5	The Data-Driven Grammar Learning method enhanced classroom interaction among students, between the student and the teacher, and the teacher and the student.	0 % 0	0 % 0	0 % 0	27 % 13	73 % 36	4.73	5	5	0.45	-1.1	-0.83
6	Classroom interaction motivates and encourages me to actively participate in the learning process.	2 % 1	2 % 1	8 % 4	29 % 14	59 % 29	4.4	5	5	0.89	-1.85	3.95
7	Exploring the language in authentic contexts was a meaningful learning experience.	2 % 1	4 % 2	10 % 5	41 % 20	43 % 21	4.18	4	5	0.93	-1.36	2.13
8	The Data-Driven Grammar Learning method created a skills-based grammar learning process that includes critical thinking, analytical, problem-solving, and collaboration skills.	4 % 2	4 % 2	8 % 4	51 % 25	33 % 16	4.04	4	4	0.98	-1.48	2.59
9	It is essential to incorporate a skills-oriented approach to grammar learning alongside a knowledge-based learning process.	2 % 1	2 % 1	12 % 6	43 % 21	41 % 20	4.18	4	4	0.88	-1.32	2.5
10	The Data-Driven Grammar Learning method promoted autonomous learning.	2 % 1	2 % 1	4 % 2	27 % 13	65 % 32	4.51	5	5	0.84	-2.3	6.29
11	Autonomous learning encouraged me to take responsibility for my own learning.	4 % 2	6 % 3	6 % 3	39 % 19	45 % 22	4.14	4	5	1.06	-1.5	1.94

12	The integration of cooperative activities, such as pair and group work, significantly increased my interest and engagement in grammar learning.	2% 1	2% 1	4% 2	33% 16	59% 29	4.45	5	5	0.84	-2.12	5.66
13	Collaborative learning contributed to my confidence-building in language learning.	2% 1	4% 2	10% 5	45% 22	39% 19	4.2	4	4	0.86	-1.42	3.01
14	Computer-assisted grammar learning increased my enthusiasm for studying grammar.	0% 0	0% 0	10% 5	14% 7	76% 37	4.65	5	5	0.66	-1.71	1.58
15	The integration of authenticity in grammar learning increased my confidence in language awareness.	6% 3	0% 0	2% 1	43% 21	49% 24	4.28	4	5	1	-2.18	5.27
16	The Data-Driven Grammar Learning method was a well-structured grammar learning method.	0% 0	2% 1	8% 4	53% 26	37% 18	4.24	4	4	0.69	-0.76	0.99
17	The leading questions provided in the DDL worksheets were helpful in developing hypotheses about the formation and use of various grammar points.	0% 0	0% 0	8% 4	25% 12	67% 33	4.59	5	5	0.64	-1.34	0.68
18	Clarifications and indications provided by the teacher along the way of working out the hypotheses were useful.	0% 0	2% 1	8% 4	47% 23	43% 21	4.3	4	4	0.71	-0.89	0.91
19	The DDL tasks were designed with clarity in mind.	2% 1	2% 1	6% 3	31% 15	59% 29	4.43	5	5	0.87	-1.98	4.74
20	The feedback provided by the teacher on the DDL tasks played a crucial role in my learning progress.	4% 2	4% 2	6% 3	35% 17	51% 25	4.24	5	5	1.03	-1.71	2.81
21	The Data-Driven Grammar Learning method is an effective approach for retaining knowledge.	4% 2	6% 3	6% 3	45% 22	39% 19	4.08	4	4	1.04	-1.45	1.98
Item	**Statement**	**Very difficult**	**Difficult**	**Of medium**	**Easy**	**Very easy**	**Mean**	**Median**	**Mode**	**Standard deviation**	**Kurtosis**	**Skewness**
22	How difficult was it to interact with a corpus system?	14% 7	19% 9	45% 22	16% 8	6% 3	2.81	3	3	1.07	-0.04	-0.28
23	How difficult was it to understand authentic language?	21% 10	20% 10	49% 25	10% 5	0% 0	2.49	3	3	0.94	-0.36	-0.84
24	How difficult was it to analyze data in KWIC format?	16% 8	33% 16	43% 21	8% 4	0% 0	2.43	3	3	0.87	-0.17	-0.66
25	How difficult was it to form hypotheses from concordance lines and leading	8%	6%	23%	45%	18%	3.59	4	4	1.12	-0.9	0.4

	questions presented in DDL worksheets?	4	3	11	22	9						
26	How difficult was it to complete DDL tasks?	8 % 4	6 % 3	31 % 15	35 % 17	20 % 10	3.53	4	4	1.14	-0.65	0.02
27	How difficult was it to learn English grammar using the DDL method?	6 % 3	8 % 4	39 % 19	31 % 15	16 % 8	3.43	3	3	1.06	-0.4	0.01

The mean is closely aligned with the median and mode for each item, which supports the reliability of the results. The standard deviation values range from 0.4 to 1.2. A low standard deviation indicates that the majority of the data points are clustered around the mean, while a high standard deviation suggests that the data is more widely spread.

Skewness and kurtosis for most items fall between -3 and 3, indicating a normal univariate distribution. However, for items 6, 10, 12, and 19, the kurtosis values exceed 3, suggesting that these datasets have heavier tails than a normal distribution - meaning there are more extreme values in the tails.

The mean scores for items 1-21 are quite high, above 4, which reflects participants' positive reactions to the statements. In contrast, the mean scores for items 22-27 are relatively low, indicating students' negative responses to those questions.

The results of the questionnaire indicated that a significant majority (94%, N=46) of students in the experimental group found the Data-Driven Learning (DDL) method to be both interesting and motivational. Additionally, 94% (N=46) of respondents reported that they were more engaged in their grammar studies through the DDL method compared to the traditional (deductive) teaching method.

Among the learners, 53% (N=26) strongly agreed, and 40% (N=19) agreed with the statement that learning grammar through the DDL method was enjoyable. Overall, 80% of participants (N=39) felt that discovering grammatical features independently was a more valuable experience than simply receiving information from teachers.

All participants (100%, N=49) responded positively regarding the DDL method's ability to enhance classroom interaction - among students, and between the student and the teacher, and the teacher and the student. 59% (N=29) strongly agreed, and 29% (N=14) agreed that classroom interaction supported and encouraged their active participation in the learning process.

Furthermore, 84% of participants (N=41) found exploring the language in authentic contexts to be a meaningful learning experience. A small percentage (10%, N=5) remained neutral on this statement, and even fewer students (6%, N=3) disagreed.

Fifty-one percent of the learners (N=25) agreed, and thirty-three percent (N=16) strongly agreed with the statement that the Data-Driven Grammar Learning (DDL) method facilitated a skills-based approach to grammar learning, which included critical thinking, analytical reasoning, problem-solving, and collaboration skills. Additionally, eighty-four percent (N=41) of the students viewed a skills-oriented grammar learning process, along with a knowledge-oriented approach, as essential. Twelve percent (N=6) were neutral on the matter, while four percent (N=2) disagreed.

Furthermore, sixty-five percent (N=32) of the respondents strongly agreed, and twenty-seven percent (N=13) agreed that the DDL method promotes autonomous learning. As a result, a significant majority (ninety-two percent, N=45) acknowledged the DDL method as a valuable tool for enhancing autonomous learning. Eighty-four percent (N=41) of participants recognized that autonomous learning encouraged them to take responsibility for their own learning. This increased autonomy, facilitated by the DDL method, has consequently led to heightened responsibility for learning among the students.

A significant majority of learners, 59% (N=29), strongly agreed, and 33% (N=16) agreed with the statement that integrating cooperative (pair and group) activities in grammar teaching increased their interest and engagement in learning grammar. Overall, this means 92% of the students felt more interested and involved in grammar learning due to the cooperative activities associated with the Data-Driven Learning (DDL) method.

Additionally, 45% (N=22) of the learners agreed, while 39% (N=19) strongly agreed that collaborative learning contributed to building their confidence in language learning. In total, a large percentage (84%, N=41) of students recognized collaborative learning as a key factor in enhancing their confidence in language acquisition.

It was also found that a significant majority of the learners (90%, N=44) expressed enthusiasm for computer-assisted grammar learning, with 76% strongly agreeing with this notion. Furthermore, 92% (N=45) of the students believed that integrating authentic materials into grammar learning helped to boost their confidence in language awareness. 90% (N=44) of the participants regarded the Data-Driven Grammar Learning method as a well-structured approach to grammar learning.

A total of 92% of the participants (N=45) responded positively regarding the inclusion of leading questions in the Data-Driven Learning (DDL) worksheets as helpful for developing hypotheses related to the formation and use of grammar points. Specifically, 67% of participants strongly agreed and 25% agreed with this statement. Additionally, 47% of the learners (N=23) agreed, and 43% (N=21) strongly agreed that receiving clarifications and indications from the teacher was useful in formulating hypotheses.

Furthermore, 90% of the students in the experimental group (N=44) assessed the DDL tasks as clearly designed. Regarding the feedback provided by the teacher on the DDL tasks, 51% of learners (N=25) strongly agreed, and 35% (N=17) agreed that it contributed positively to their learning progress. An impressive 84% of the learners (N=41) recognized the Data-Driven Grammar Learning method as effective for knowledge retention.

The second part of the questionnaire focused on the difficulties students might have encountered while using the Data-Driven Learning method. It was found that 45% of learners (N=22) perceived interaction with a corpus as of medium difficulty, while 33% (N=16) rated it as very difficult or difficult, and 22% (N=11) found it easy or very easy. Thus, the majority (45%) viewed corpus interaction as moderately challenging, whereas a significant portion (33%) considered it difficult.

Regarding the understanding of authentic language, 21% of students (N=10) found it very difficult to comprehend, and 20% (N=10) considered it difficult. In contrast, 49% (N=24) reported that understanding authentic language was of medium difficulty, while 10% (N=5) found it easy. Hence, most students (49%) perceived the comprehension of authentic language as moderately challenging, although a substantial proportion (41%) thought it was difficult to understand.

A survey of learners revealed varying levels of difficulty when analyzing data in KWIC format. Specifically, 16% of the learners (N=8) found it very difficult, 33% (N=16) considered it difficult, 43% (N=21) reported it as medium difficulty, and 8% (N=4) found it easy. Overall, a significant majority of students - 49% - perceived analyzing data in KWIC format as hard, while another large portion, 43%, rated it as medium difficulty.

When it came to forming hypotheses based on concordance lines and leading questions presented in DDL worksheets, 14% of learners (N=7) found this task very difficult or difficult, while 23% (N=11) experienced medium difficulty. A notable 63% (N=31) reported that they found it easy or very easy. These findings suggest that a considerable number of students did not encounter difficulties in developing hypotheses within the context of DDL worksheets.

Regarding the accomplishment of DDL tasks, 14% of students (N=7) rated them as very difficult or difficult, 31% (N=15) experienced medium difficulty, and 55% (N=27) found the tasks easy or very easy. In summary, 55% of participants found DDL tasks to be easy or very easy, while the next largest group (31%) rated them as medium difficulty.

In terms of learning English grammar through the DDL method, 14% of students (N=7) found it very difficult or difficult, 39% (N=19) rated it as medium difficulty, and 47% (N=23) found it easy or very easy. Overall, the largest group (47%) considered the DDL method to be easy or very easy, while the second largest group (39%) reported it as medium difficulty.

When asked about their preferences for continuing to learn EFL grammar, 45% of students expressed a desire to continue using the DDL method, 49% preferred a combination of the DDL method and the traditional teaching approach, and 7% showed a preference for the traditional teaching method alone.

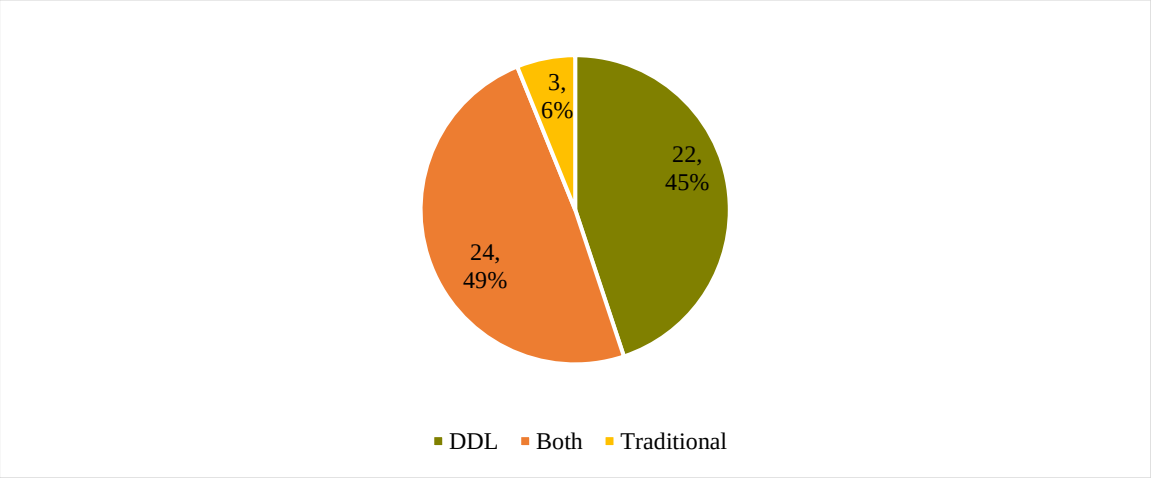

Figure 16. Students' preferences for EFL grammar learning.

The survey results indicated that a slightly higher number of students preferred learning EFL grammar through a combination of both the Data-Driven Learning (DDL) method and the traditional method, compared to those who preferred only the DDL method. Overall, the preference for these combined methods surpassed that for the traditional approach alone.

Based on the majority of responses, it can be inferred that students in both experimental groups (the Data-Driven Learning group and the mixed-methods group) found Data-Driven Grammar Learning to be an engaging and motivational approach to grammar instruction. Most students reported feeling more engaged in their grammar studies using the Data-Driven Learning method compared to the traditional, deductive approach. They recognized the experience of discovering grammatical features independently as a valuable part of their learning process.

The Data-Driven Grammar Learning method enhanced classroom interaction, encouraging students to actively participate. Exploring language in authentic contexts provided them with a meaningful learning experience. Participants viewed Data-Driven Learning as a skills-based approach to grammar learning and considered it significant.

Students indicated that this method promoted autonomous learning, empowering them to take responsibility for their educational journey. They also noted that cooperative work in grammar learning heightened their interest and involvement. Collaborative learning efforts helped increase their confidence in language acquisition.

Additionally, students mentioned that the computer-assisted grammar learning method boosted their motivation for studying grammar. Engaging with authentic language materials contributed to their awareness and confidence. They recognized the Data-Driven Grammar Learning method as well-structured and appreciated the leading questions provided in the DDL worksheets, which guided them in developing hypotheses.

Participants assessed the DDL tasks as clearly designed and beneficial. Based on their feedback, teacher-provided guidance on DDL tasks significantly contributed to their learning progress. Ultimately, students found the Data-Driven Grammar Learning method effective for knowledge retention.

Furthermore, the majority of students perceived interacting with a corpus as having medium difficulty, which was similar to their experience with comprehending authentic language. However, nearly half of the students found analyzing data in KWIC (Keyword in Context) format to be challenging. Interestingly, students did not report significant difficulties in developing hypotheses using DDL worksheets, and they rated the completion of DDL tasks as easy. Specifically, 47% of the students in the experimental group found Data-Driven Learning of EFL grammar easy, while 39% rated it as medium difficulty. Additionally, 49% of the students expressed a preference for continuing their EFL grammar studies through both the DDL method and the traditional method of teaching, while 45% preferred solely the DDL method.

The findings from the second questionnaire are summarized below.

When asked about their preferred method of grammar instruction, 94% of the students in the experimental groups (N=46) opted for the DDL method or a combination of the DDL method with traditional teaching. Only 6% of the learners (N=3) favored the traditional method of grammar instruction. This response aligns with the results from the final question of the first questionnaire mentioned earlier.

The students' responses about their preference for traditional methods of grammar learning over Data-Driven Learning (DDL) are presented below. The students were allowed to answer the question in Georgian, their native language, and all three students expressed their opinions in Georgian. Translations of their responses are provided below.

Student 1: *I prefer when my teacher provides me with grammatical rules and explanations. This approach to grammar learning feels easier for me. I can apply the rules to different examples, which makes the process enjoyable and straightforward. I can complete a lot of grammatical exercises without difficulty. This method of learning is more comfortable for me, and I believe it allows me to learn faster. I have been studying grammar this way for a long time, and I would like to continue with the traditional method.*

Student 2: *I found the DDL method to be technically challenging. I struggled to master the part-of-speech tags and their various uses. The corpus interface and its tools were also difficult for me to navigate. Although I successfully completed the tasks in the DDL worksheets prepared by the teacher, working within the corpus itself was challenging. I support technology-enhanced language teaching; however, I find corpus-based teaching to be very difficult, and I believe it is unnecessary to include technology in grammar instruction.*

Student 3: *I did not like the principle of the DDL method, which involves reaching conclusions based on examples from a corpus and developing rules independently. I found this process challenging; group work was helpful, as I would have struggled to do it on my own. Additionally, I found working with incomplete sentences in the corpus to be very inconvenient.*

In summary, the students highlighted several reasons for preferring the traditional method of English as a Foreign Language (EFL) grammar teaching over the DDL method, or a combination of the DDL method with the traditional approach:

1. The traditional method is perceived as an easier way to learn.
2. The traditional method is seen as a quicker way to learn.
3. Students are accustomed to the old (deductive teaching) method.
4. There are difficulties related to technology and the technical skills required for the DDL method.
5. The inductive nature of Data-Driven Learning involves reasoning and deducing rules, which can be challenging.
6. Data is often presented in the form of incomplete sentences.

To summarize, 6% of the learners in the experimental groups preferred the conventional method of teaching EFL grammar due to the challenges posed by certain features of the DDL method, the ease of explicit instruction, and their familiarity with the traditional learning method.

The findings of this research align with previous studies that indicated EFL university students have positive attitudes toward the Data-Driven Grammar Learning approach.

While the results of this study are partially consistent with those reported by Girgin (2011), there are notable differences. In Girgin's study, university students found learning English grammar through corpus-based activities to be useful. They did not perceive grammar learning through these activities as more difficult than learning via a coursebook, nor did they find corpus-based activities to be more boring. However, some outcomes from Girgin's research differ from the current study's findings. Specifically, the majority of students in Girgin's study indicated that their participation in corpus-based grammar activities was not particularly active, primarily because it was their first experience with corpus-based teaching. Additionally, many students disagreed with the notion that corpus-based activities enhanced their confidence in learning English grammar. Interviews revealed that students had little enthusiasm for engaging with corpus-based activities. Overall, the students' attitudes toward these grammar activities were rated as neutral, which distinguishes Girgin's study from the present research.

The findings of the present research align with those of Nugraha, Miftakh, and Wachyudi (2016), who discovered that university students felt positively about learning grammar through the Data-Driven Learning (DDL) approach. The students enjoyed the DDL course and expressed a desire to learn more grammar using this method.

Additionally, they found DDL worksheets, grammar explanations, follow-up activities, and teacher feedback to be beneficial.

Furthermore, the results of this research are consistent with the findings of Elsherbini and Ali (2017). In their study, tertiary education students reported that using a corpus was both easy and effective for learning English grammar. These students indicated that their confidence was boosted through corpus-based learning and expressed a willingness to continue using corpora for grammar studies in the future. They also agreed that teachers should incorporate corpora into English as a Foreign Language (EFL) instruction. However, interviews with some students revealed challenges associated with corpus usage. Specifically, some students found corpus-based learning to be time-consuming, struggled with interpreting concordance results and cut-off sentences, and faced technical difficulties, such as registration issues. These challenges are somewhat similar to those identified by students in the current study.

The findings of the current study align partially with the research conducted by Yanto and Nugraha (2017), which explored university students' perceptions of corpus-aided discovery learning in English grammar. Their study revealed that students had varying feelings about the use of corpora in grammar learning. While many students found corpus-based activities challenging, all participants acknowledged that corpus-based instruction was useful. More than half of the students indicated that learning grammar through corpus-informed activities was more difficult than using a traditional coursebook. However, these same students found corpus-informed activities to be less tedious than coursebook tasks. Most students reported their level of participation in corpus-informed activities as active or somewhat active. A significant majority also believed that these activities helped improve their grammar skills and boosted their confidence in learning English grammar. Despite these positive outcomes, students had differing opinions regarding the preference for corpus application over coursebook use. Their attitudes towards recommending corpus use in grammar learning were similarly divided. Overall, the students involved in Yanto and Nugraha's research expressed ambiguous and sometimes conflicting views regarding various aspects of corpus-based teaching. Their perspectives on corpus-aided discovery learning in English grammar can be characterized as neutral, rather than distinctly positive or negative.

The findings of the current study align with those of Abdul-Ameer (2019). The University EFL students who participated in Abdul-Ameer’s research reported a positive attitude toward Data-Driven Learning (DDL) of English grammar. A large majority of these students found learning grammar through the DDL approach to be interesting and enjoyable. They believed they learned more effectively by exploring the language and developing the rules independently. The students considered immediate feedback from their teacher to be valuable. They also assessed DDL homework activities as beneficial for retaining knowledge. Furthermore, a significant majority of the students expressed a preference for integrating DDL with traditional teaching methods, while a small percentage showed a willingness to continue with conventional grammar instruction.

C. Conclusion to Study 2

The results of the study indicated that students in the experimental groups had positive attitudes toward various aspects of the Data-Driven Grammar Learning method. Most responses highlighted that Data-Driven Learning of grammar was both interesting and motivational, making it an enjoyable approach to learning. Students found it to be more engaging than the traditional grammar teaching method. The investigative nature of this method was viewed as a more valuable experience compared to explicit teaching. Classroom interactions, enhanced by the DDL method, encouraged learners to actively participate in the learning process. Students recognized the exploration of authentic language as a valuable component of their learning experience. Data-Driven Learning of Grammar was acknowledged as a skills-based approach, which students considered essential alongside knowledge-based learning. Additionally, responses indicated that the DDL method promoted autonomous learning, leading to increased responsibility for their education among students. The results confirmed that cooperative activities improved motivation and engagement. Furthermore, collaborative learning was seen as a facilitator for building confidence in language acquisition. Computer-assisted grammar learning was found to foster enthusiasm for learning grammar. The incorporation of authentic materials in grammar instruction positively impacted students’ self-confidence regarding language awareness. Students rated the Data-Driven Learning (DDL) tasks as well-organized and clearly designed. The guiding questions presented in the DDL worksheets, along with the teacher’s directions for developing hypotheses and providing feedback on the DDL tasks, were recognized as helpful components of the method. Furthermore, the DDL approach was considered effective for knowledge retention. A significant majority of

students in the experimental groups (94%) expressed a desire to continue experiencing Data-Driven Learning for grammar in the future, whether purely or in conjunction with the traditional method of grammar teaching. Only a small percentage of students (6%) preferred the traditional method of grammar instruction, citing challenges related to the technological and technical skills required for the Data-Driven Grammar Learning method, the ease of explicit instruction compared to inductive reasoning, the data presentation format, often consisting of incomplete sentences, and their familiarity with the conventional learning method.

7. CONCLUSIONS AND RECOMMENDATIONS

The extensive literature reviewed in the monograph reveals that corpus-based Data-Driven Learning (DDL) - whether conducted through paper-based, computer-based, or a combination of both approaches - offers unique opportunities for language learners. Students can engage with authentic language, observe the usage and frequency of various language patterns, identify language trends, and conduct linguistic investigations.

The analysis highlights several benefits of Data-Driven Learning. Firstly, exposure to real-life language enhances language awareness among students. DDL fosters active learning, as students actively examine different facets of language. Additionally, it encourages students to become autonomous, self-regulated learners. This method creates a student-centered environment in which students take on the role of researchers while teachers act as facilitators of learning. The process of inquiry and inference aligns with constructivist learning principles. Moreover, Data-Driven Learning is characterized as an interactive and enjoyable educational experience, providing students with opportunities to develop essential lifelong skills such as analytical thinking, problem-solving, and critical reasoning.

Although DDL is often perceived as suitable for high-proficiency language learners, numerous studies reviewed in this monograph demonstrate its effectiveness for students at the elementary language level as well. While Data-Driven Learning presents certain challenges, such as the complexity and authenticity of language data, scholars have proposed various strategies to address these issues.

The model of Data-Driven Learning of Grammar developed in this monograph includes the following steps: *Observation and Deduction, Presentation of Findings, Clues and Fine-Tuning, Corpus Interaction, Practice and Feedback,* and *Knowledge Assessment.*

A one-year experimental study aimed at measuring the development of undergraduate students' grammar skills through Data-Driven Learning of Grammar involved 71 freshmen from a higher education institution in Georgia. The study comprised three groups: a) a traditional teaching group (control group), b) a Data-Driven Learning (DDL) group (experimental group 1), and c) a mixed-method group (DDL combined with traditional instruction) (experimental group 2). The results of the experiment, which included a pre-test, two midterm tests, a post-test, and a delayed test, revealed that the control group had a gain of 16.91, while experimental group 1 achieved a gain of 31.21, and experimental group 2 had a gain of 33.96. Paired-samples t-tests demonstrated the statistical significance of the differences in achievements between the pairs: the significance level was $p = 0.000$, which is less than 0.05 in all three comparisons (control group vs. experimental group 1, control group vs. experimental group 2, and experimental group 1 vs. experimental group 2).

The mixed-methods group slightly outperformed the pure Data-Driven Learning (DDL) group. Overall, the experimental study demonstrated that Data-Driven Learning of Grammar - whether through pure DDL or in combination with the traditional teaching method - was more effective in improving undergraduate EFL students' grammar skills compared to traditional grammar instruction. Additionally, Data-Driven Learning of Grammar proved to be effective in terms of knowledge retention. It's important to note that while the mixed-methods group achieved slightly better overall results than the pure DDL group, the pure DDL method led to longer-lasting learning outcomes. Students in the DDL group exhibited better knowledge retention than those in the mixed-methods group. Therefore, the pure DDL method is not less effective, and both methods are recommended for use.

A survey aimed at exploring the attitudes of students in experimental groups toward Data-Driven Learning of Grammar, which included 49 students from both experimental groups, revealed that the learners had positive perceptions of various aspects of this method. The majority of the students recognized Data-Driven Learning of Grammar as:

- A motivational and enjoyable process for learning grammar.

- More engaging than the traditional (deductive) teaching method.
- A more valuable experience than receiving ready-made knowledge from the teacher.
- A facilitator of enhanced classroom interaction.
- A meaningful learning experience that offers opportunities to explore the language in authentic contexts.
- A skills-based approach that involves critical thinking, analytical skills, problem-solving skills, collaboration skills, and more, alongside knowledge-oriented learning.
- A method that encourages autonomous learning, leading to increased responsibility for their education.
- An incentive for greater motivation and engagement due to the cooperative activities involved.
- A promoter of confidence in language learning through collaborative experiences.
- A boost to enthusiasm for learning grammar, particularly because of its integration with technology.
- A well-structured approach to grammar learning.
- An effective method for knowledge retention.

The vast majority of learners expressed a desire to continue learning English as a Foreign Language (EFL) grammar through either Data-Driven Learning (DDL) or a mixed-method approach that combines DDL with the traditional teaching method. The small percentage of students (6%) who preferred the conventional approach to grammar learning cited several reasons for their choice, including:

- Difficulties related to technology and the technical skills required for the DDL method.
- The ease of receiving explicit instruction compared to the challenges of inductive reasoning.
- The presentation of data in formats that included incomplete sentences.
- A familiarity with the conventional teaching method.

Overall, the Data-Driven Grammar Learning approach is a student-centered, constructivist, authentic, language-based, skills-oriented, and autonomous method for learning grammar. The research presented in this monograph demonstrates the effectiveness of Data-Driven Learning of Grammar in enhancing English as a Foreign Language (EFL) students' grammar skills and fostering motivational and active grammar learning.

Based on the findings of this monograph, the following recommendations can be made:

1. It is highly recommended to implement corpus-based grammar teaching, either in its pure form or combined with the traditional approach, in the university EFL context. This method helps students connect with real-life language and promotes the development of higher-order thinking and lifelong learning skills.
2. Considering various scholarly perspectives and the results of numerous studies, the implementation of Data-Driven Learning of Grammar is recommended at all levels of English language teaching.
3. It is advisable for teachers to use the model of Data-Driven Learning of Grammar proposed in this monograph.
4. Conducting a larger-scale experimental study with a broader participant sample, particularly including students at proficiency levels lower or higher than B1, is recommended. This could lead to the development of different models and strategies for Data-Driven Learning of Grammar.
5. A post-experimental study investigating students' attitudes and the difficulties they encounter during the learning process is also suggested.
6. Additionally, a qualitative study could be conducted to explore the challenges and issues teachers face in the process of data-driven grammar instruction.

These recommendations aim to enhance the overall effectiveness and adaptability of Data-Driven Grammar Learning in educational contexts.

References

Abdul-Ameer, M.A. (2019). Corpora in the EFL classroom: Exploring the effects of Data Driven Learning (DDL) on Iraqi EFL freshmen's grammatical development. *Journal of Basra Researches for Human Sciences*, *44*(4), 21-33.

Akinci, M., & Yildiz, S. (2017). Effectiveness of corpus consultation in teaching verb + noun collocations to advanced ELT students. *Eurasian Journal of Applied Linguistics*, *3*(1), 91-108.

Al-Gamal, A.A.M., & Ali, E.A.M. (2019). Corpus-based method in language learning and teaching. *International Journal of Research and Analytical Reviews (IJRAR)*, *6*(2), 473-476.

Anani Sarab, M.R., & Kardoust, A. (2014). Concordance-based data-driven learning activities and learning English phrasal verbs in EFL classrooms. *Issues in Language Teaching*, *3*(1), 112-89.

Ashouri, S., Arjmandi, M., & Rahimi, R. (2014). The impact of corpus-based collocation instruction on Iranian EFL learners' collocation learning. *Universal Journal of Educational Research*, *2*(6), 470-479.

Barabadi, E., & Khajavi, Y. (2017). The effect of data-driven approach to teaching vocabulary on Iranian students' learning of English vocabulary. *Cogent Education*, *4*(1), 1-13.

Bedmar, M.D. (2006). Data-driven learning and awareness-raising: An effective tandem to improve grammar in written composition? *International Journal of Emerging Technologies in Learning (iJET)*, *1*(3), 1-11.

Bennett, G.R. (2010). *Using Corpora in the Language Learning Classroom: Corpus Linguistics for Teachers* (Vol. 10). Ann Arbor, MI: University of Michigan Press.

Biber, D., & Reppen, R. (2002). What does frequency have to do with grammar teaching? *Studies in Second Language Acquisition*, *24*, 199-208.

Binkai, J. (2012). An empirical study on corpus-driven English vocabulary learning in China. *English Language Teaching*, *5*(4), 131-137.

Blass, L., Iannuzzi, S., Savage, A., Gordon, D., & Reppen, R. (2013). *Grammar and Beyond* Level 3 Student's book. Cambridge: Cambridge University Press.

Boontam, P., & Phoocharoensil, S. (2018). Effectiveness of English Preposition Learning through Data-Driven Learning (DDL). *3L: Language, Linguistics, Literature*, *24*(3), 125-141.

Boulton, A. (2008). Looking (for) empirical evidence of data-driven learning at lower levels. In B. Lewandowska-Tomaszczyk (Ed.), *Corpus Linguistics, Computer Tools, and Applications: State of the Art* (volume 17) (pp. 581-598). Frankfurt: Peter Lang.

Boulton, A. (2009). Data-driven learning: Reasonable fears and rational reassurance. *Indian Journal of Applied Linguistics*, *35*(1), 81-106.

Boulton, A. (2010). Data-driven learning: on paper, in practice. In T. Harris & M. Moreno Jaén (eds.) *Corpus Linguistics in Language Teaching* (pp. 17-52). Bern: Peter Lang.

Boulton, A. (2012). Hands-on/hands-off: Alternative approaches to data-driven learning. In J. Thomas & A. Boulton (Eds.), *Input, Process and Product: Developments in Teaching and Language Corpora* (pp. 152-168). Masaryk: Masaryk University Press.

Boulton, A. (2016). Integrating corpus tools and techniques in ESP courses. *ASp. Ia revue du GERAS*, *69*, 113-137.

Braun, S. (2007). Integrating corpus work into secondary education: From data-driven learning to needs-driven corpora. *ReCALL*, *19*, 307-328.

Brown, M.H. (2017). Using the sentence corpus of remedial English to introduce Data-Driven Learning tasks. *Kanda Academic Review*, *1*(1), 1-14.

Carter, R., & McCarthy, M. (1995). Grammar and the spoken language. *Applied Linguistics*, *16*(2), 141-158.

Çelik, S. (2011). Developing collocational competence through web based concordance activities. *Novitas-ROYAL (Research on Youth and Language)*, *5(*2), 273-286.

Chan, T.P., & Liou, H.C. (2005). Effects of web-based concordancing instruction on EFL students' learning of verb–noun collocations. *Computer Assisted Language Learning*, *18*(3), 231-251.

Chirobocea, O. (2017). The good and the bad of the corpus-based approach (or data-driven learning) to ESP teaching. *"Mircea cel Batran" Naval Academy Scientific Bulletin*, *20*(1), 364-371.

Chujo, K., & Oghigian, K. (2007). Discovering grammar basics with parallel concordancing in the beginner-level EFL classroom. Retrieved from http://hanamizuki2010.sakura.ne.jp/public_html/data/2007%20Discovering%20grammar%20basic.pdf

Chujo, K., & Oghigian, K. (2008). A DDL approach to learning noun and verb phrases in the beginner level EFL classroom. In *Proceedings of 8th Teaching and Language Corpora (TaLC) Conference* (pp. 65-71). Lisbon, Portugal: ISLA Lisboa.

Chujo, K., Anthony, L., Oghigian, K., & Uchibori, A. (2012). Paper-based, computer-based, and combined data-driven learning using a web-based concordancer. *Language Education in Asia*, *3*(2), 132-145.

Chujo, K., Anthony, L., Oghigian, K., & Yokota, K. (2013). Teaching remedial grammar through data-driven learning using AntPConc. *Taiwan International ESP Journal*, *5*(2), 65-90.

Chujo, K., Kobayashi, Y., Mizumoto, A., & Oghigian, K. (2016). Exploring the effectiveness of combined web-based corpus tools for beginner EFL DDL. *Linguistics and Literature Studies*, *4*(4), 262-274.

Chujo, K., Utiyama, M., & Miura, S. (2006). Using a Japanese-English parallel corpus for teaching English vocabulary to beginning-level students. *English Corpus Studies*, *13*, 153-172.

Dash, N.S. (2011). Corpus-Based English Language Teaching: A New Method. In *Proceedings of the Orientation-Cum-Workshop Programme on Recent Trends in Language Teaching* (pp. 19-36). Kolkata, India: David Hare Training College, Institute of Advanced Studies in Education.

Dazdarevic, S., Zoranic, A. L., & Fijuljanin, F. (2015). Benefits of corpus-based approach to language teaching. *Balkan Distance Education Network - BADEN Newsletter*, *7*, 1-5. Retrieved from https://www.academia.edu/16213074/Benefits_of_corpus_based_approach_to_language_teaching?auto=download

De Jesus, A.D.C. & Carrilo, J.L.Q. (2014). Advantages of using corpora to teach English. Retrieved from http://cenedic.ucol.mx/fieel/pdf/1.pdf

Demir, C. (2017). Lexical collocations in English: A comparative study of native and non-native scholars of English. *Dil ve Dilbilimi Çalışmaları Dergisi*, *13*(1), 75-87.

Elsherbini, S.A.H., & 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.

Flowerdew, L. (2009). Applying corpus linguistics to pedagogy: A critical evaluation. *International Journal of Corpus Linguistics*, *14*(3), 393-417.

Frankenberg-Garcia, A. (2012). Raising teachers' awareness of corpora. *Language Teaching*, *45*(4), 475-489.

Geluso, J., & Yamaguchi, A. (2014). Discovering formulaic language through data-driven learning: Student attitudes and efficacy. *ReCALL*, *26*(2), 225-242.

Gilquin, G., & Granger, S. (2010). How can data-driven learning be used in language teaching. In A. O'Keeffe & M. McCarthy (Eds.). *The Routledge Handbook of Corpus Linguistics* (pp. 359-370). New York: Routledge.

Girgin, U. (2011). *Corpus-Based Activities at Lower Levels of EFL Proficiency: The Effectiveness of Using Concordance Lines on Grammar Learning* (Master's Thesis). Ankara: Bilkent University.

Girgin, U. (2019). The effectiveness of using corpus-based activities on the learning of some phrasal-prepositional verbs. *Turkish Online Journal of Educational Technology-TOJET*, *18*(1), 118-125.

GNC (2022). Corpus list. *The Georgian National Corpus.* Retrieved from http://gnc.gov.ge/gnc/page

Hadley, G. (2002). An introduction to data-driven learning. *RELC Journal, 33*(2), 99-124.

Hanafiyeh, M., & Keshi, A. K. (2013). The effect of corpus-based instruction and thesaurus-based teaching on Iranians EFL learners' grammatical knowledge. *Journal of Basic and Applied Scientific Research, 3*(2), 167-179.

Hardie, A. (2012). CQPweb - combining power, flexibility and usability in a corpus analysis tool. *International Journal of Corpus Linguistics, 17*(3), 380–409. Available at https://cqpweb.lancs.ac.uk/ .

Hoffmann, S. & Evert, S. (2018). *The CQP-edition of BNCweb* (Version 4). Available at http://corpora.lancs.ac.uk/BNCweb/ .

Huang, L.S. (2008). Using guided, corpus-aided discovery to generate active learning. *English Teaching Forum, 46*(4), 20-27.

Huang, L.S. (2011). Corpus-aided language learning. *ELT Journal, 65*(4), 481-484.

Huang, L.S. (2018). Corpus-based instruction. *The TESOL Encyclopedia of English Language Teaching, 1,* 1-11.

Hunston, S. (2002). *Corpora in Applied Linguistics.* Cambridge: Cambridge University Press.

Jafarpour, A., Hashemian, M., & Alipour, S. (2013). A corpus-based approach toward teaching collocation of synonyms. *Theory and Practice in Language Studies, 3*(1), 51-60.

Jalilifar, A., Mehrabi, K., & Mousavinia, S. R. (2014). The effect of concordance enriched instruction on the vocabulary learning and retention of Iranian EFL learners. *Procedia-Social and Behavioral Sciences, 98,* 742-746.

Johansson, S. (2009). Some thoughts on corpora and second-language acquisition. In K. Aijmer (ed.), *Corpora and Language Teaching* (pp. 33-44). Amsterdam: John Benjamins.

Johns, T. (1990). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *CALL Austria, 10,* 14–34.

Johns, T.F. (1991). Should you be persuaded: Two samples of data-driven learning materials. *ELR Journal, 4,* 1–16.

Jones, C., & Waller, D. (2015). *Corpus Linguistics for Grammar: A Guide for Research*. London: Routledge.

Kayaoglu, M. (2013). The use of corpus for close synonyms. *Journal of Language and Linguistic Studies, 9*(1), 128-144.

Kazuko (2014). Using a corpus in an EFL classroom to develop depth of vocabulary knowledge. *The Tsuru University Revirew, 79,* 133-146.

Kenchiashvili N. (2020). History of Corpus Linguistics as a new interdisciplinary field in Georgia (corpora and webcorpora). *Language and Culture, 23,* 38-50.

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., ... & Suchomel, V. (2014). The Sketch Engine: ten years on. *Lexicography, 1*(1), 7-36. Available at available at https://www.sketchengine.eu .

Kim, H. (2019). The perception of teachers and learners towards an exploratory corpus-based grammar instruction in a Korean EFL primary school context. *Primary English Education, 25*(1), 123-152.

Koosha, M., & Jafarpour, A.A. (2006). Data-driven learning and teaching collocation of prepositions: The case of Iranian EFL adult learners. *Asian EFL Journal, 8*(4), 192-209.

Krashen, S. (1985). *The Input Hypothesis: Issues and Implications.* Oxford: Pergamon Press.

Li, S. (2017). Using corpora to develop learners' collocational competence. *Language Learning & Technology, 21*(3), 153-171.

Lin, M.H., & Lee, J.Y. (2015). Data-driven learning: Changing the teaching of grammar in EFL classes. *Elt Journal, 69*(3), 264-274.

Liu, D., & Jiang, P. (2009). Using a corpus-based lexico-grammatical approach to grammar instruction in EFL and ESL contexts. *The Modern Language Journal, 93*(1), 61-78.

Mann, M., & Taylore-Knowles, S. (2016a). *LASER B1* (third edition). London: Macmillan Education.

Mann, M., & Taylore-Knowles, S. (2016b). *LASER B1+* (third edition). London: Macmillan Education.

Mansoor, M. S., & Salman, Y. M. (2013). Collocation, colligation and semantic prosody. *Buhuth Mustaqbaliya Scientific Periodical Journal*, *4*(2), 1-34.

Marinov, S. (2013). Training ESP students in corpus use-challenges of using corpus-based exercises with students of non-philological studies. *Teaching English with Technology, 13*(4), 49-76.

McCarthy, M., McCarten, J., & Sandiford, H. (2014). *Touchstone* Level 4 Student's Book (second edition). Cambridge: Cambridge University Press.

Moehkardi, R.R.D. (2002). *Grammatical and lexical English collocations: Some possible problems to Indonesian learners of English*. *Humaniora, 14*(1), 53-62.

Nugraha, S.I., Miftakh, F., & Wachyudi, K. (2016). Teaching grammar through data-driven learning (DDL) approach. *Advances in Social Science, Education and Humanities Research (ASSEHR), 82,* 300-303.

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

O'Sullivan, I. (2007). Enhancing a process-oriented approach to literacy and language learning: The role of corpus consultation literacy. *ReCALL: The Journal of EUROCALL*, *19*(3), 269-286.

Phoocharoensil, S. (2012). Language corpora for EFL teachers: An exploration of English grammar through concordance lines. *Procedia-Social and Behavioral Sciences, 64*, 507-514.

Qiong, H. (2017). Corpora and English teaching: Retrospect and prospect. *International Journal of Scientific Research and Innovative Technology, 4*(7), 65-73.

Rahimi, M., & Momeni, G. (2012). The effect of teaching collocations on English language proficiency. *Procedia-Social and Behavioral Sciences, 31*, 37-42.

Rapti, N. (2010). *A Study of Classroom Concordancing in the Greek Context: Data-driven Grammar Teaching and Adolescent EFL Learners* (Doctoral Dissertation). Nottingham, United Kingdom: University of Nottingham.

Reppen, R. (2010). *Using Corpora in the Language Classroom*. Cambridge: Cambridge University Press.

Römer, U. (2010). Corpora and language teaching. In A. Lüdeling & M. Kytö. (Eds.). *Corpus Linguistics* (Vol. 1) (pp. 112-130). Berlin: Walter de Gruyter.

Sahillioglu, F., Şahinkayasi, Y., & Şahinkayasi, H. (2011). Effectiveness of two corpus-based approaches on Turkish student's learning English grammar. In Z. Genc (Ed.), *5th International Computer and Instructional Technologies Symposium Proceedings Book* (pp. 261-269). Elazig: Fırat University Printing Office.

Sahragard, R., Kushki, A., & Ansaripour, E. (2013). The application of corpora in teaching grammar: The case of English relative clause. *Journal of Pan-Pacific Association of Applied Linguistics, 17*(1), 79-93.

Schmidt, R.W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*(2), 129-158.

Schmidt, R.W. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, *13*, 206-226.

Smart, J. (2014). The role of guided induction in paper-based data-driven learning. *ReCALL: the Journal of EUROCALL*, *26*(2), 184-201.

Smith, S. (2011). Learner construction of corpora for general English in Taiwan. *Computer Assisted Language Learning*, *24*(4), 291-316.

Sun, X., & Hu, G. (2020). Direct and indirect data-driven learning: An experimental study of hedging in an EFL writing class. *Language Teaching Research, 1*, 1-29.

Sun, Y.C., & Wang, L.Y. (2003). Concordancers in the EFL classroom: Cognitive approaches and collocation difficulty. *Computer Assisted Language Learning*, *16*(1), 83-94.

Takahashi, S., & Fujiwara, Y. (2016). Effects of inductive learning based on Data-Driven Learning at elementary schools in Japan. *JES Journal*, *16*(1), 84-99.

Tandashvili M., & Purtskhvanadze, Z. (2014). *Digital Humanities - Kartvelology and the Challenges of the 21st Century*. Frankfurt: Frankfurt University Press.

Ucar, S., & Yükselir, C. (2015). The effect of corpus-based activities on verb-noun collocations in EFL Classes. *Turkish Online Journal of Educational Technology (TOJET)*, *14*(2), 195-205.

Unaldi, I., Bardakci, M., Akpinar, K.D., & Dolas, F. (2013). A comparison of contextualized, decontextualized and corpus-informed vocabulary instruction: A quasi experimental study. *Journal of Language and Literature Education*, *8*, 78-95.

Varley, S. (2009). I'll just look that up in the concordancer: integrating corpus consultation into the language learning environment. *Computer Assisted Language Learning*, *22*(2), 133-152.

Vyatkina, N. (2016). Data-driven learning of collocations: Learner performance, proficiency, and perceptions. *Language Learning & Technology, 20(3),* 159–179.

Vyatkina, N. (2020). Corpora as open educational resources for language teaching. *Foreign Language Annals*, *53*(2), 359-370.

Wang, X. (2018). Effect of a corpus-based grammar teaching method in the Chinese EFL environment. *Advances in Social Science, Education and Humanities Research (ASSEHR), 182,* 294-297.

Yanto, E.S., & Nugraha, S.I. (2017). The implementation of corpus-aided discovery learning in English grammar pedagogy. *Journal of ELT Research: The Academic Journal of Studies in English Language Teaching and Learning*, *2*(2), 66-83.

Yilmaz, M. (2017). The effect of data-driven learning on EFL students' acquisition of lexico-grammatical patterns in EFL writing. *Eurasian Journal of Applied Linguistics*, *3*(2), 75-88.

Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing*, *13*(4), 257-283.

Yoon, H., & Jo, J. (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), 96-117.

Yunus, K. (2014). The impact of Data-Driven Learning instruction on Malaysian law undergraduates' colligational competence. *Kajian Malaysia*, *32*(1), 79-109.

Appendixes

Appendix 1. Questionnaire 1 (Investigating the Experimental Groups Students' Attitudes towards Data-Driven Learning of Grammar) for Study 1

Questionnaire 1 for survey research *Exploring Experimental Groups Students' Attitudes towards Data-Driven Learning of Grammar*

Dear students,

The purpose of the survey is to learn about your attitudes towards Data-Driven Grammar Learning method. Completion of the following questionnaire will require no more than 15 minutes. Your responses will remain completely anonymous; confidentiality will also be observed. The results will be used only in International Black Sea University as a part of my PhD dissertation and will not be disclosed to any other third party.

კითხვარი კვლევისთვის ექსპერიმენტული ჯგუფების სტუდენტების დამოკიდებულებების შესწავლა კორპუს-მონაცემებზე დაფუძნებული გრამატიკის სწავლების მეთოდის მიმართ

ძვირფასო სტუდენტებო,

გამოკითხვის მიზანია გაარკვიოს თქვენი დამოკიდებულებები კორპუს-მონაცემებზე დაფუძნებული გრამატიკის სწავლების მეთოდის მიმართ. კითხვარის შევსებას დასჭირდება არაუმეტეს 15 წუთისა. თქვენი პასუხები დარჩება ანონიმური; ასევე დაცული იქნება კონფიდენციალურობა. შედეგები გამოყენებული იქნება მხოლოდ როგორც ჩემი სადოქტორო დისერტაციის ნაწილი შავი ზღვის საერთაშორისო უნივერსიტეტში. თქვენი მონაწილეობა ძალიან ღირებულია.

The asterisk (*) indicates questions that are mandatory to respond.

1. The Data-Driven Grammar Learning method is an interesting and motivational approach to learning grammar. კორპუს მონაცემებზე დაფუძნებული გრამატიკის სწავლების მეთოდი (The Data-Driven Grammar Learning/DDL method) არის საინტერესო და მოტივაციური მეთოდი გრამატიკის შესასწავლად.*
   - ☐ Strongly disagree სრულიად არ ვეთანხმები
   - ☐ Disagree არ ვეთანხმები
   - ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
   - ☐ Agree ვეთანხმები
   - ☐ Strongly agree სრულიად ვეთანხმები
2. I was more engaged in my grammar study through the DDL method than the traditional (deductive) teaching method. DDL მეთოდის საშუალებით უფრო მეტად ვიყავი ჩართული გრამატიკის სწავლის პროცესში, ვიდრე უწინ ტრადიციული (დედუქციური) სწავლების მეთოდისას.*
   - ☐ Strongly disagree სრულიად არ ვეთანხმები
   - ☐ Disagree არ ვეთანხმები
   - ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
   - ☐ Agree ვეთანხმები
   - ☐ Strongly agree სრულიად ვეთანხმები

3. Learning grammar through the DDL method was enjoyable. გრამატიკის სწავლა DDL მეთოდით იყო სახალისო.*
    - ☐ Strongly disagree სრულიად არ ვეთანხმები
    - ☐ Disagree არ ვეთანხმები
    - ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
    - ☐ Agree ვეთანხმები
    - ☐ Strongly agree სრულიად ვეთანხმები
4. Discovering grammatical features on my own was a more valuable experience than simply receiving ready-made knowledge from my teacher. გრამატიკული მახასიათებლების დამოუკიდებლად აღმოჩენა უფრო ღირებული გამოცდილება იყო ჩემთვის, ვიდრე მასწავლებლისგან მზა ცოდნის მიღება.*
    - ☐ Strongly disagree სრულიად არ ვეთანხმები
    - ☐ Disagree არ ვეთანხმები
    - ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
    - ☐ Agree ვეთანხმები
    - ☐ Strongly agree სრულიად ვეთანხმები
5. The Data-Driven Grammar Learning method enhanced classroom interaction among students, between the student and the teacher, and the teacher and the student. DDL მეთოდმა გააძლიერა საკლასო ინტერაქცია სტუდენტებს შორის, სტუდენტსა და მასწავლებელს შორის, და მასწავლებელსა და სტუდენტს შორის.*
    - ☐ Strongly disagree სრულიად არ ვეთანხმები
    - ☐ Disagree არ ვეთანხმები
    - ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
    - ☐ Agree ვეთანხმები
    - ☐ Strongly agree სრულიად ვეთანხმები
6. Classroom interaction motivates and encourages me to actively participate in the learning process. საკლასო ინტერაქცია მეხმარება და სტიმულს მაძლევს რომ უფრო აქტიურად ჩავერთო სასწავლო პროცესში.*
    - ☐ Strongly disagree სრულიად არ ვეთანხმები
    - ☐ Disagree არ ვეთანხმები
    - ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
    - ☐ Agree ვეთანხმები
    - ☐ Strongly agree სრულიად ვეთანხმები
7. Exploring the language in authentic contexts was a meaningful learning experience. ენის შესწავლა ავთენტურ კონტექსტში იყო მნიშვნელოვანი სასწავლო გამოცდილება. *
    - ☐ Strongly disagree სრულიად არ ვეთანხმები
    - ☐ Disagree არ ვეთანხმები
    - ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
    - ☐ Agree ვეთანხმები
    - ☐ Strongly agree სრულიად ვეთანხმები
8. The Data-Driven Grammar Learning method created a skills-based grammar learning process that includes critical thinking, analytical, problem-solving, and collaboration skills. DDL მეთოდმა წარმოშვა უნარებზე დაფუძნებული გრამატიკის სწავლის პროცესი, რომელიც მოიცავს კრიტიკულ აზროვნებას, ანალიტიკურ, პრობლემების გადაჭრისა და თანამშრომლობის უნარებს. *
    - ☐ Strongly disagree სრულიად არ ვეთანხმები
    - ☐ Disagree არ ვეთანხმები
    - ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
    - ☐ Agree ვეთანხმები
    - ☐ Strongly agree სრულიად ვეთანხმები

9. It is essential to incorporate a skills-oriented approach to grammar learning alongside a knowledge-based learning process. აუცილებელია უნარებზე ორიენტირებული მიდგომის ინტეგრირება გრამატიკის სწავლისადმი ცოდნაზე-დაფუძნებულ სწავლასთან ერთად.*
    - ☐ Strongly disagree სრულიად არ ვეთანხმები
    - ☐ Disagree არ ვეთანხმები
    - ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
    - ☐ Agree ვეთანხმები
    - ☐ Strongly agree სრულიად ვეთანხმები
10. The Data-Driven Grammar Learning method promoted autonomous learning. DDL მეთოდმა ხელი შეუწყო ავტონომიურ სწავლას. *
    - ☐ Strongly disagree სრულიად არ ვეთანხმები
    - ☐ Disagree არ ვეთანხმები
    - ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
    - ☐ Agree ვეთანხმები
    - ☐ Strongly agree სრულიად ვეთანხმები
11. Autonomous learning encouraged me to take responsibility for my own learning. ავტონომიურმა სწავლამ მიბიძგა, რომ პასუხისმგებლობა ამეღო საკუთარ სწავლაზე. *
    - ☐ Strongly disagree სრულიად არ ვეთანხმები
    - ☐ Disagree არ ვეთანხმები
    - ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
    - ☐ Agree ვეთანხმები
    - ☐ Strongly agree სრულიად ვეთანხმები
12. The integration of cooperative activities, such as pair and group work, significantly increased my interest and engagement in grammar learning. თანამშრომლობითი აქტივობების, როგორიცაა წყვილებში და ჯგუფურ მუშაობა, ინტეგრაციამ მნიშვნელოვნად გაზარდა ჩემი ინტერესი და ჩართულობა გრამატიკის სწავლაში. *
    - ☐ Strongly disagree სრულიად არ ვეთანხმები
    - ☐ Disagree არ ვეთანხმები
    - ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
    - ☐ Agree ვეთანხმები
    - ☐ Strongly agree სრულიად ვეთანხმები
13. Collaborative learning contributed to my confidence-building in language learning. თანამშრომლობითმა სწავლამ ჩემი თავდაჯერებულობის გაზრდას შეუწყო ხელი ენის შესწავლის პროცესში. *
    - ☐ Strongly disagree სრულიად არ ვეთანხმები
    - ☐ Disagree არ ვეთანხმები
    - ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
    - ☐ Agree ვეთანხმები
    - ☐ Strongly agree სრულიად ვეთანხმები
14. Computer-assisted grammar learning increased my enthusiasm for studying grammar. კომპიუტერის დახმარებით გრამატიკის სწავლამ გაზარდა ჩემი ენთუზიაზმი გრამატიკის შესწავლის მიმართ.*
    - ☐ Strongly disagree სრულიად არ ვეთანხმები
    - ☐ Disagree არ ვეთანხმები
    - ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
    - ☐ Agree ვეთანხმები
    - ☐ Strongly agree სრულიად ვეთანხმები
15. The integration of authenticity in grammar learning increased my confidence in language awareness. გრამატიკის შესწავლაში ავთენტურობის ინტეგრაციამ გაზარდა ჩემი თავდაჯერებულობა ენობრივი ცნობიერების მიმართ.*
    - ☐ Strongly disagree სრულიად არ ვეთანხმები

- ☐ Disagree არ ვეთანხმები
- ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
- ☐ Agree ვეთანხმები
- ☐ Strongly agree სრულიად ვეთანხმები

16. The Data-Driven Grammar Learning method was a well-structured grammar learning method. DDL მეთოდი იყო კარგად სტრუქტურირებული გრამატიკის სწავლის მეთოდი. *
    - ☐ Strongly disagree სრულიად არ ვეთანხმები
    - ☐ Disagree არ ვეთანხმები
    - ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
    - ☐ Agree ვეთანხმები
    - ☐ Strongly agree სრულიად ვეთანხმები
17. The leading questions provided in the DDL worksheets were helpful in developing hypotheses about the formation and use of various grammar points. DDL სამუშაო ფურცლებში მოცემული წამყვანი კითხვები სასარგებლო იყო სხვადასხვა გრამატიკული ერთეული ფორმირებისა და გამოყენების შესახებ ჰიპოთეზების შემუშავებაში. *
    - ☐ Strongly disagree სრულიად არ ვეთანხმები
    - ☐ Disagree არ ვეთანხმები
    - ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
    - ☐ Agree ვეთანხმები
    - ☐ Strongly agree სრულიად ვეთანხმები
18. Clarifications and indications provided by the teacher along the way of working out the hypotheses were useful. მასწავლებლის მიერ მოწოდებული განმარტებები და მითითებები იყო სასარგებლო ჰიპოთეზების შემუშავების გზაზე.*
    - ☐ Strongly disagree სრულიად არ ვეთანხმები
    - ☐ Disagree არ ვეთანხმები
    - ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
    - ☐ Agree ვეთანხმები
    - ☐ Strongly agree სრულიად ვეთანხმები
19. The DDL tasks were designed with clarity in mind. DDL დავალებები გასაგებად იყო დიზაინირებული.*
    - ☐ Strongly disagree სრულიად არ ვეთანხმები
    - ☐ Disagree არ ვეთანხმები
    - ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
    - ☐ Agree ვეთანხმები
    - ☐ Strongly agree სრულიად ვეთანხმები
20. The feedback provided by the teacher on the DDL tasks played a crucial role in my learning progress. მასწავლებლის მიერ DDL დავალებებთან დაკავშირებით მოწოდებულმა უკუკავშირმა გადამწყვეტი როლი ითამაშა ჩემს სწავლის პროგრესში. *
    - ☐ Strongly disagree სრულიად არ ვეთანხმები
    - ☐ Disagree არ ვეთანხმები
    - ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
    - ☐ Agree ვეთანხმები
    - ☐ Strongly agree სრულიად ვეთანხმები
21. The Data-Driven Grammar Learning method is an effective approach for retaining knowledge. DDL მეთოდი ეფექტური მეთოდია მიღებული ცოდნის შესანარჩუნებლად. *
    - ☐ Strongly disagree სრულიად არ ვეთანხმები
    - ☐ Disagree არ ვეთანხმები
    - ☐ Neither disagree nor agree არც ვეთანმხები, არც არ ვეთანმხები
    - ☐ Agree ვეთანხმები
    - ☐ Strongly agree სრულიად ვეთანხმები

22. How difficult was it to interact with a corpus system? რამდენად რთული იყო კორპუსის სისტემასთან ურთიერთობა? *
    - ☐ Very difficult ძალიან რთული
    - ☐ Difficult რთული
    - ☐ Of medium difficulty საშუალო სირთულის
    - ☐ Easy მარტივი
    - ☐ Very easy ძალიან მარტივი
23. How difficult was it to understand authentic language? რამდენად რთული იყო ავთენტური ენის გაგება? *
    - ☐ Very difficult ძალიან რთული
    - ☐ Difficult რთული
    - ☐ Of medium difficulty საშუალო სირთულის
    - ☐ Easy მარტივი
    - ☐ Very easy ძალიან მარტივი
24. How difficult was it to analyze data in KWIC format? რამდენად რთული იყო კორპუს მონაცემთა KWIC ფორმატში ანალიზი? *
    - ☐ Very difficult ძალიან რთული
    - ☐ Difficult რთული
    - ☐ Of medium difficulty საშუალო სირთულის
    - ☐ Easy მარტივი
    - ☐ Very easy ძალიან მარტივი
25. How difficult was it to form hypotheses from concordance lines and leading questions presented in DDL worksheets? რამდენად რთული იყო ჰიპოთეზების ჩამოყალიბება DDL სამუშაო ფურცლებში წარმოდგენილი კონკორდანს მაგალითებზე და წამყვანი კითხვებზე დაყრდნობით? *
    - ☐ Very difficult ძალიან რთული
    - ☐ Difficult რთული
    - ☐ Of medium difficulty საშუალო სირთულის
    - ☐ Easy მარტივი
    - ☐ Very easy ძალიან მარტივი
26. How difficult was it to complete DDL tasks? რამდენად რთული იყო DDL დავალებების შესრულება? *
    - ☐ Very difficult ძალიან რთული
    - ☐ Difficult რთული
    - ☐ Of medium difficulty საშუალო სირთულის
    - ☐ Easy მარტივი
    - ☐ Very easy ძალიან მარტივი
27. How difficult was it to learn English grammar using the DDL method? რამდენად რთული იყო ინგლისური გრამატიკის შესწავლა DDL მეთოდით? *
    - ☐ Very difficult ძალიან რთული
    - ☐ Difficult რთული
    - ☐ Of medium difficulty საშუალო სირთულის
    - ☐ Easy მარტივი
    - ☐ Very easy ძალიან მარტივი
28. I would like to continue learning EFL grammar through ვისურვებდი ინგლისური, როგორც უცხოური ენის გრამატიკის სწავლის გაგრძელებას *
    - ☐ the DDL method DDL მეთოდის გამოყენებით
    - ☐ both the DDL method and the traditional method of teaching ორივე მეთოდის გამოყენებით, როგორც DDL მეთოდის, ასევე ტრადიციული სწავლების მეთოდის გამოყენებით
    - ☐ the traditional method of teaching სწავლების ტრადიციული მეთოდის გამოყენებით

Questionnaire 2 (Investigating the Reasons for the Experimental Groups Students' Preference for the Traditional Method of Grammar Teaching) for Study 2

Questionnaire 2 for survey research *Exploring Experimental Groups Students' Attitudes towards Data-Driven Learning of Grammar*

Dear students,

The purpose of the survey is to find out your preference for the EFL grammar teaching method and the reasons for your preference for the traditional method of grammar teaching. Completion of the following questionnaire will require no more than 15 minutes. Your responses will remain completely anonymous; confidentiality will also be observed. The results will be used only in International Black Sea University as a part of my PhD dissertation and will not be disclosed to any other third party.

კითხვარი მინი-კვლევისთვის *სტუდენტების უპირატესობების გამოკვლევა ინგლისური როგორც უცხოური ენის გრამატიკის სწავლების მეთოდების მიმართ*

ძვირფასო სტუდენტებო,

გამოკითხვის მიზანია დაადგინოს ინგლისური როგორც უცხოური ენის გრამატიკის სწავლების რომელ მეთოდს ანიჭებთ უპირატესობას. კითხვარის შევსებას დასჭირდება არაუმეტეს 15 წუთისა. თქვენი პასუხები დარჩება ანონიმური; ასევე დაცული იქნება კონფიდენციალურობა. შედეგები გამოყენებული იქნება მხოლოდ როგორც ჩემი სადოქტორო დისერტაციის ნაწილი შავი ზღვის საერთაშორისო უნივერსიტეტში. თქვენი მონაწილეობა ძალიან ღირებულია.

The asterisk (*) indicates questions that are mandatory to respond.

1. I would like to continue learning EFL grammar using ვისურვებდი ინგლისური, როგორც უცხოური ენის გრამატიკის სწავლის გაგრძელებას *
   - ☐ either the DDL method or a combination of the DDL method and the traditional teaching method DDL მეთოდის გამოყენებით ან DDL მეთოდის ტრადიციულ მეთოდთან შერწმის გამოყენებით
   - ☐ the traditional teaching method სწავლების ტრადიციული მეთოდის გამოყენებით Going to the Question 2

2. Name the reasons for preferring the traditional method of teaching grammar to the Data-Driven Grammar Learning method. (You can answer the question in Georgian) დაასახელეთ მიზეზები თუ რატომ ანიჭებთ უპირატესობას გრამატიკის სწავლების ტრადიციულ მეთოდს კორპუს-მონაცემებზე დაფუძნებული გრამატიკის სწავლების მეთოდთან შედარებით (შეკითხვაზე პასუხი შეიძლება გასცეთ ქართულ ენაზე) *