LEARN Journal: Language Education and Acquisition Research Network
(ISSN: 2630-0672 (Print) | ISSN: 2672-9431 (Online)
Volume: 15, No: 2, July – December 2022
Language Institute, Thammasat University
https://so04.tci-thaijo.org/index.php/LEARN/index

Effects of Using Online Corpus and Online Dictionary as Data-driven Learning on Students' Grammar Mastery

Akhmad Fauzan[a], Yazid Basthomi[b*], Francisca Maria Ivone[c]

[a] a_fauzan@edu.upr.ac.id, English Language Teaching, Faculty of Letters, Universitas Negeri Malang, Indonesia
[b] ybasthomi@um.ac.id, English Language Teaching, Faculty of Letters, Universitas Negeri Malang, Indonesia
[c] francisca.maria.fs@um.ac.id, English Language Teaching, Faculty of Letters, Universitas Negeri Malang, Indonesia
[*] Corresponding author, ybasthomi@um.ac.id

APA Citation:
Fauzan, A., Basthomi, Y., & Ivone, F.M. (2022). Effects of using online corpus and online dictionary as data-driven learning on students' grammar mastery. *LEARN Journal: Language Education and Acquisition Research Network, 15*(2), 679-704.

Received 07/12/2021

Received in revised form 05/06/2022

Accepted 20/06/2022)

Keywords
online corpus, online dictionary, data-driven learning, English grammar

Abstract

This study compares the effects of using online corpus and online dictionary as Data-Driven Learning (DDL) on EFL students' grammar mastery and their perceptions after learning with the two media in grammar instruction. The experiment with a counterbalanced design was employed with two replications, and 65 students in two intact classes participated. In the first replication, 33 students of the first group learned grammar with an online corpus, while 32 students of the second group learned using an online dictionary. In the second replication, the first group of students learned grammar with an online dictionary, while the second group learned with an online corpus. After each replication, the students took a grammar test and filled in a questionnaire. The grammar tests after learning with an online corpus were compared to the results of the grammar tests after learning with an online dictionary. Then the

	students' responses on both media were also compared. The statistical evidence comparing grammar test results reveals no statistical difference in students' grammar mastery after learning with an online corpus and an online dictionary. The same result is obtained when comparing their perceptions after learning with both media.

Introduction

Data-Driven learning (DDL) has been extensively used in language learning (Boulton & Cobb, 2017; Lee et al., 2019) using language data derived from corpus or dictionary applications. A number of studies pertinent to corpora have been conducted in grammar classrooms and the results show that language learners gain benefits from the use of corpora. Corpus use is a highly recommendable approach in university level EFL grammar classrooms because intermediate to advanced level learners may benefit from DDL approach (Lin, 2016). Authentic texts provided by a corpus allow language learners to see examples of contextual grammar use to increase their awareness of syntactic patterns and make fewer grammatical mistakes (Girgin, 2019; Lin & Lee, 2015; Türkmen & Aydin, 2016). As a result, the use of corpora can promote the teaching of grammar in EFL contexts.

Besides corpora as a language data tool, dictionaries also have advantageous features to language learners, particularly in vocabulary instruction (Y. Chen, 2016; Chiu & Liu, 2013; Dwaik, 2015) as they offer examples of sentences that enable the learners to see the patterns of colligation as models (Nurmukhamedov, 2012; Tananuraksakul, 2015). Dictionaries have been used in reading and writing classrooms due to their features of showing definitions, synonyms, or examples (Chiu & Liu, 2013; Dilenschneider, 2017; Yoon, 2016). However, there is a lack of evidence on the effects of dictionary, particularly the online one, in grammar instruction. In a well-developed online dictionary website, for example *YourDictionary*, there are many sentences to show as a result of a word search which is unlike the most print or electronic dictionaries offer, as they show only one or two sentences. This online dictionary displays authentic example sentences that can be a source of learning for DDL because these language data can be used to identify patterns or rules of grammatical items in sentences.

In previous studies, corpora and dictionaries were interchangeably used in vocabulary instruction to equip language learners with information on word usage, word meaning, and collocation information (Gao, 2011; Karras, 2016; Lai & Chen, 2015; Tsai, 2019; Yoon, 2016). The two media have also been compared in vocabulary teaching and learning in EFL classrooms due to their potential to provide rich and authentic English language data (Frankenberg-Garcia, 2014; Z. Huang, 2014; Mueller & Jacobsen, 2016). Corpora and dictionaries are comparable for vocabulary instruction due to the features they offer. Most web-based corpora and dictionaries collect language data from various sources and transform the data to be examples of sentences. Besides showing the results of the searched word in the form of sentences, corpora also have other features, such as collocations and synonyms. On the other hand, online dictionaries offer searching word definitions, thesaurus, and examples.

Though the two media have been interchangeably used and compared in vocabulary instruction, there has been relatively small evidence on studies comparing the effects of using the two media on EFL students' grammar mastery. The two media are potential for Data-Driven Learning (DDL) in grammar instruction due to their capability to provide digitalized language data in English sentences. Therefore, this study compares a corpus website *Sketch Engine for Language Learning* (SkELL) to a web-based dictionary, *YourDictionary*, by counterbalancing them and thus finds out their effects on EFL students' grammar mastery. The current study also seeks the students' perceptions on the use of both media.

Literature Review

Data-Driven Learning (DDL)

DDL is characterized as an approach in which the language learner is a researcher whose learning is influenced by linguistic data (Johns, 1991). The proposed idea is that the role of learners in DDL is like researchers who investigate particular language patterns, discover the rules, and develop their analytical skills. The underlying idea of DDL is similar to the concept of the inductive method in general that teachers convert the classrooms into student-centered by asking them to identify patterns, find rules from examples, and finally use the patterns in practice. For example, during grammar instruction, the process in the inductive

method begins with learners analyzing examples to discover the rules of grammar. This activity encourages them to recognize how grammatical items function (Fleming, 2018; Pawlak, 2018). The examples used in this method can be a sample of languages, such as a text or a set of sentences containing the targeted grammatical items (Basturkmen, 2018). As a result, language learners can improve their grammatical accuracy because they can access many examples. Hence, they elicit the grammatical rules from these examples. Then, the information is processed and stored in their long-term memory as they work things out for themselves, rather than simply being given the principle or rule (Alzu'bi, 2015; Nunan, 1998).

In order to implement DDL in language classrooms, teachers can provide pre-prepared samples of sentences for learners to analyze or offer them corpus access so that they can search the data on their own. Teachers' roles include facilitating the learning process by preparing learners to utilize corpus and draw conclusions from the corpus data and motivating them to work through vast amounts of corpus data (Chambers, 2010; Flowerdew, 2012; Sripicharn, 2010). Language learners practice their analytical skills through DDL when they see the language data. They can then use their skills to analyze other sentences. Language learners are trained to observe and think about what grammatical rules work in authentic sentences. However, they are also given the opportunities to generalize the grammar rules independently (Lin & Lee, 2015). As a result, they can develop grammatical awareness, actively strive to avoid grammar errors in the future, and show significant improvements in grammar tests after a teacher-led guided DDL induction (Mizumoto & Chujo, 2016; Moon & Oh, 2017).

DDL also significantly improves language learners' overall learning motivation and self-efficacy in grammar, and they respond positively towards the implementation of DDL in learning grammar. They admit that their awareness of lexical items is increased, and they can solve lexico-grammatical problems encountered during the writing and eventually develop their grammatical consciousness (Aşık et al., 2016; Boontam & Phoocharoensil, 2018; M. Chen & Flowerdew, 2018). DDL treatments also appear to transform them into active learners, allow them to learn new words by referring to multiple samples of how a target lexical item is used differently in various contexts, and motivate them to correct errors of word choice and word form (Crosthwaite, 2017; Lee et al., 2019; Lin & Lee, 2015; Luo, 2016).

Nowadays, DDL has used well-documented language data from corpora as they are available in several forms. The first category of the corpus is available and established programs provided by well-known institutions. These general corpora are derived from spoken and written sources (fiction, popular magazines, newspapers, and academic texts), i.e., *Corpus of Contemporary American English* (COCA), *British National Corpus* (BNC), and *Michigan Corpus of Academic Spoken English* (MICASE) which contain millions of language data in the form of collocations and example sentences (Bardovi-Harlig et al., 2017; Chang, 2014; Daskalovska, 2015). The second category of the corpus is the specifically developed corpora. It can be topic-specific corpora developed based on the needs for a specific discipline or a specialized, multi-disciplinary corpus annotated with genre conventions for input enhancement (Cotos et al., 2017; Laosrirattanachai & Ruangjaroon, 2021; Poole, 2016).

Online Corpus

Corpora have been implemented in English classrooms in some English as foreign language countries. For example, in Japan, despite the lack of readily available authentic language materials in the classroom, corpora provide language learners with an advantage by providing a large number of example sentences provided per search and determining how particular parts of speech collocate (Hirata & Hirata, 2019; Mueller & Jacobsen, 2016). Furthermore, since the Grammar Translation Method is prevalent in Taiwan's grammar classrooms, corpora have transformed students into active learners, and they welcome and favor corpora as a fresh, engaging, and flexible method (Lin, 2016; Lin & Lee, 2015). Furthermore, corpora have been used in Korean universities as reference sources and are valued for their direct help in academic writing (Chang, 2014).

Corpora potentially help language learners improve their vocabulary by allowing them to correct the use of words they understand but frequently misuse and hypothesize how to make the natural language (Frankenberg-Garcia, 2012; Gordani, 2013). Besides, they can make a broader range of appropriate collocations because they can investigate language patterns in corpora and use some of these patterns to accurately resemble the patterns of professionals in the field (Ackerley, 2017; Friginal, 2013). Furthermore, corpus-based activities allow language learners to

spend some time analyzing and interpreting the information, implying that there is a depth of processing, allowing them to learn and recall the knowledge more effectively (Daskalovska, 2015).

Language learners can enhance the quality of their writing by accessing corpora since it improves their written expression, particularly during revision tasks, and corpus referencing has proven to be a helpful experience for them (Quinn, 2014; Tono et al., 2014). In the writing process, language learners use corpus tools to retrieve sample sentences that contain the keyword of their choice before attempting to correct their errors (Lai & Chen, 2015; Larsen-Walker, 2017). In terms of the writing products, their writings include a higher rate of accuracy or naturalness and increased use of academic collocations and fixed phraseological items (Li, 2017). In addition, by accessing a considerable number of language data, they are exposed to various examples and can identify how particular words work in contexts.

Some studies have also been conducted concerning implementing a corpus in grammar instruction. Authentic texts from online concordancers promote grammar teaching in EFL learning and teaching processes (Türkmen & Aydin, 2016). Language learners are not only satisfied with the exploration of language data provided by corpora, but they also start to understand some grammar points that they had previously struggled to grasp through traditional methods, such as memorization (Phoocharoensil, 2012). As a result, corpus-based activities are excellent for teaching grammar in two dimensions: form and use because language learners can recognize and comprehend the form of phrasal-prepositional verbs through the activities (Girgin, 2019).

When corpora are utilized in grammar instruction, some variables contribute to language learners' grammatical progress. Corpus learning allows learners to identify and understand the structure of words and their use in context, allowing them to reuse them to construct new sentences. Because the learning activities focus on self-discovery, exposure to target structures in authentic and genuine situations increases language learners' analytical skills. Additionally, it is considered that including relevant and contextualized materials assisted them in learning and increased retention of the grammar learned (Girgin, 2019; Lin & Lee, 2015).

Online Dictionary

In EFL classrooms, dictionary use has been investigated on the production and retention of collocations. Y. Chen (2016) claims that using a dictionary (the *Oxford Advanced Learner's English-Chinese Dictionary*) significantly contributes to the increase of language learners' productive collocation knowledge. Chiu and Liu (2013) confirm that most of the participants in their study agree that electronic dictionaries are more convenient to use, primarily because they can quickly retrieve the meanings of unfamiliar words. They also believe that using a dictionary can help them remember words because they must look up the word meanings and thus make connections with new vocabulary items.

Dwaik (2015) has found that language learners who use online dictionaries have a higher reading proficiency average than those who use electronic and printed dictionaries. Though the dictionaries are used mainly for reading comprehension and translation, the learners' overall language proficiency improved significantly after being exposed to authentic language found in electronic dictionaries. Moreover, they prefer electronic and online dictionaries over printed dictionaries because online dictionaries offer more authentic entries with fewer vocabulary items and more complex grammar. As an online dictionary is used in a writing classroom, Tananuraksakul (2015) reveals that *Cambridge Dictionaries Online* improves learners' English during writing assignments by using the patterns and the example sentences as writing models. The learners in the study positively view the use of the online dictionary as they agree that it is beneficial, practical, and trendy. Thus, they are more motivated to write better in English. Furthermore, the online dictionary assists them in studying English autonomously to some extent because they moderately use it. In addition, the learners are enthusiastic about continuing to use the *Cambridge Dictionaries Online* to enhance their writing skills.

Overall, online dictionaries are preferred over other forms of dictionaries because they can quickly obtain word definitions and provide sample sentences. This form of dictionary allows language learners to learn without interruption, making it easier for them to recall words. They also promote learning autonomy since they value the usage of dictionaries in their studies and are eager to utilize them while learning a foreign language. In writing lessons, online dictionaries' example sentences have

also been used as model sentences. However, the use of dictionaries in grammar learning has not been thoroughly studied.

Methodology

Procedure

Experimental research with a counterbalanced design was employed with the flow of research implementation is described as Replication 1 and 2. In Replication 1, the first group of students learned the present perfect tense with an online corpus for four meetings, while the second group learned the same tense using an online dictionary for four meetings. In Replication 2, the first group of students learned the passive voice with an online dictionary for four meetings, while the second group learned the same construction with an online corpus for four meetings. After each replication, the students took a grammar test and filled in a questionnaire. Due to pandemic of Covid-19, the teaching and learning activities were conducted fully online by employing *Zoom Cloud Meetings* software.

In each meeting of grammar instruction, Data-Driven Learning (DDL) was conducted following the steps. In the first group, the students were told that they would learn present perfect tense in active sentences. The first step of Data-driven Learning (DDL) was searching language data of present perfect sentences. The teacher gave one key word *seen* and asked the students to use browsers on their laptops or smartphones to access the *SkELL* website and use the first feature of *SkELL*, which is *Examples*. The next step was identifying sentences. As the search results showed sentences, the teacher asked them to identify sentences containing any form of *has*/*have* that appears before the word *seen*. They had to copy the sentences on the website and send it to the teacher privately via *Chat* in *Zoom Cloud Meetings*. While the students were making this identification, the teacher also copied some sentences from *SkELL* to *Microsoft Word* for the analysis later on. After that, he led them to the next step of DDL, which was analyzing the sentences. The teacher shared the screen displaying the sentences in *Microsoft Word* to the students. The teacher invited them to analyze each sentence by breaking it down into syntactic structures (subject, predicate, object, complement) as well as coloring it based on the structures or parts of speech. The next

step of DDL was classifying the findings. The teacher asked the students to group the sentences based on the auxiliary verb *has*/*have* or adverbial markers. As they finished grouping the sentences, he led them to the next step which was generating rules. Here the teacher invited the students to tell the class about the rules that underlay the grouping they previously made. It might be based on the subject of the sentence (so that *has* or *have* was used) or the adverb of time (*since* or *for*). Next, the teacher moved on to other new key words to search and repeated the steps above. At the end of the meeting, the teacher proceeded to the last step of DDL which was sentence production. He asked the students to develop one sentence of present perfect for each of the searched words they had discussed. During the sentence development, the students might use the second menu of *SkELL* (*Word Sketch*).

The activities elaborated above were also done to the second group, yet they used *YourDictionary* website. The key words that the students should search in the online dictionary were different than the words used in the *SkELL* group.

The online corpus used in the current study was the *Sketch Engine for Language Learning* (SkELL). It is a web-based online corpus and can be accessed via https://skell.sketchengine.eu/#home. It is a simple tool for checking how a particular phrase or a word is used by speakers of English. The language data are useful for English language learning. There are three main menus in *SkELL*, namely *Examples*, *Word Sketch*, and *Synonyms*, but only the first two features were utilized in this study. The *Examples* section displays samples of sentences in which the searched word or phrase is used in context. It lists up to 40 examples in complete sentences. The *Word Sketch* feature is a summary of the most typical collocations divided into logical categories available for every content word, i.e., noun, adjective, adverb, or verb. It can show, for instance, the subjects for a verb, the objects for a verb, the adjectives for a noun, or the modifiers for a noun.

The online dictionary used in the present study was *YourDictionary*. It is a web-based online dictionary and can be accessed via https://www.yourdictionary.com/. This monolingual dictionary (English-English) offers a simple layout and a menu. Of all the options in the menu offered by this online dictionary, the *Dictionary* and *Sentences* features were utilized in this study. The *Dictionary* feature shows the definitions and parts of speech of words, while the *Sentences* feature can be used to

find complete sentences using a word. The genre of the language data is general and not specified for English language learning.

Participants

The current study was conducted at the English Language Teaching department of one public university in Indonesia and involved first-semester students as the participants. There were 65 students in total, and they were divided into two groups comprising 33 students in the first group and 32 students in the second group. At the beginning of the semester, the two groups of students were given a test developed by the department to measure their initial English language knowledge. Based on the results of the test, a statistical analysis was employed to examine the difference of both groups. The result of the Independent T-Test shows that the significance level between the groups is 1.00 and this indicates that $p>.05$. Therefore, the two groups are not statistically different and have the chance to be taken as the sample for this study.

Instruments

Since there were two replications in this research, there were two grammar tests conducted, one for each replication. In Replication 1, the test measured students' grammar mastery of the present perfect tense in active sentences. In this study, it was named Grammar Test 1. In Replication 2, the test measured students' grammar mastery of passive sentences, and it was named Grammar Test 2. Both Grammar Tests 1 and 2 contained two parts, namely multiple-choice questions (15 items) and sentence development (5 items). The multiple-choice questions measured students' knowledge of grammar, and the sentence development test items measured performance in using the grammar.

In the present study, questionnaires were used to collect data related to the students' perceptions towards the implementation of an online corpus and an online dictionary as Data-Driven Learning (DDL) in grammar instruction. There were two questionnaires, one for each learning format. The questionnaires contained statements related to students' perceptions of the effectiveness of activities in learning grammar with an online corpus and an online dictionary. These statements were developed based on two factors of psychometric scale to measure

learners' perceived preferences and benefits of DDL developed and validated by Mizumoto et al. (2016). The two factors (Clarity and Autonomy) were then transformed into three dimensions, namely language data, knowledge and ability, and learning experience. The statements about students' attitudes towards language data confirm their perceptions on the variety of the examples of the sentences and the authenticity of examples of sentences. In terms of knowledge and ability, the students gave their perceptions on how the media helped them identify sentence patterns, practice their sentence analysis skills, and write better English sentences. Finally, the last dimension, which is about learning experience, confirmed their perceptions of whether or not they found the media fun, effective, motivating, and helped them remember grammar rules, as well as their future plan of using the media for other grammar materials.

The questionnaires were developed by adopting the Likert Scale. There were ten items in form of statement and there were four options in every statement (Strongly Disagree, Disagree, Agree, and Strongly Agree), due to the concern that certain respondents might use the middle category (neither agree nor disagree, not sure, or neutral) to avoid making a real choice (Dörnyei & Taguchi, 2010).

Data Collection and Analysis

The data were collected in the fall semester of the 2020/2021 academic year, which runs from September 2020 to January 2021. In order to collect the data, the researchers underwent several stages. The first stage of data collection was done when the two groups completed the first intervention in Replication 1. The data were in the form of the results of the grammar test and questionnaire after the students learned using an online corpus in the first group, and the results of grammar test and questionnaire after they learned using an online dictionary in the second group. The second stage of data collection was conducted when the two groups concluded the second intervention in Replication 2. The data were in the form of the results of a grammar test and questionnaire after the students learned using an online dictionary in the first group and the results of the grammar test and questionnaire after they learned using an online corpus in the second group.

Since the current study investigated the effects of an online corpus and an online dictionary, then the collected data were rearranged based on the two media used in the treatment. The first data set was the results of learning grammar with the online corpus (data of grammar tests from the first group in Replication 1 and from the second group in Replication 2). The second data set was the results of learning grammar with the online dictionary (data of grammar tests from the second group in Replication 1 and from the first group in Replication 2). The third data set was the students' perceptions after learning with the online corpus (data of questionnaires from the first group in Replication 1 and from the second group in Replication 2). The fourth data set was the students' perceptions after learning with the online dictionary (data of questionnaires from the second group in Replication 1 and from the first group in Replication 2). Table 1 shows the organized data for analysis.

Table 1

Data for Analysis

Analysis	Data	Number of Data
Grammar Mastery	Online Corpus (Group 1 of Replication 1 + Group 2 of Replication 2)	65
	Online Dictionary (Group 2 of Replication 1 + Group 1 of Replication 2)	65
Perceptions	Online Corpus (Group 1 of Replication 1 + Group 2 of Replication 2)	65
	Online Dictionary (Group 2 of Replication 1 + Group 1 of Replication 2)	65

The analysis conducted to the first two data groups firstly tested the statistical assumptions (normality test and homogeneity test) and then compared the data of the students' grammar mastery based on the scores after learning grammar using an online corpus and an online dictionary. The second analysis involving the last two data groups compared the students' perceptions after learning grammar using an online corpus and an online dictionary.

Results

Students' Grammar Mastery

In this study, the normality and homogeneity tests were conducted to fulfill the statistical assumptions. These tests were conducted as the prerequisite in hypotheses testing for the grammar test data. There are two sets of data generated from the grammar tests: data after learning with the online corpus (N=65, Mean=76.53) and data after learning with the online dictionary (N=65, Mean=74.61). The normality test was done, and the significance level or the *p*-value from Kolmogorov-Smirnov was taken for consideration in this study due to the number of the data that are more than 50 (N=65). Table 2 shows the test of normality of the students' grammar mastery. Since the significance levels or the *p*-values for the two sets of data are .00 and .00 respectively, the $p<.05$ means that the data from the grammar tests after learning with the online corpus and online dictionary are not normally distributed.

Table 2

Normality Test of Students' Grammar Mastery

Media	Kolmogorov-Smirnov		
	Statistic	df	Sig.
Online Corpus	.16	65	.00
Online Dictionary	.15	65	.00

The second test was the homogeneity test. The Levene's Test employed in this study was used to test if two sets of data of grammar mastery have equal variances. Table 3 shows the results of the test. The significance level of the homogeneity is .79. Because $p>.05$, the data from the grammar tests after learning with the online corpus and online dictionary are equal or homogenous. As the two assumption tests were completed and the results show that the data are not normally distributed but homogenous, the statistical assumptions are not fulfilled.

Table 3

Homogeneity Test of Students' Grammar Mastery

Levene Statistic	Sig.

.06	.79

Since the statistical assumptions are not fulfilled, the nonparametric testing Mann-Whitney U-Test was performed to analyze the data sets further. The test is equivalent to the Independent Sample T-Test when comparing differences between two independent groups. The result of the calculation is summarized in Table 4. The statistical evidence comparing grammar tests results reveals no statistical difference in students' grammar mastery after learning using online corpus or online dictionary (U = 1980.00, p = .53).

Table 4

Mann-Whitney U Test of Students' Grammar Mastery

Tests	Students' Grammar Mastery
Mann-Whitney U	1980.00
Wilcoxon W	4125.00
Z	-.62
Asymp. Sig. (2-tailed)	.53

Students' Perceptions

In order to test the statistical hypothesis of the questionnaires, the statistical assumptions are not necessarily done because the questionnaires have ordinal data. The Mann-Whitney U-Test was employed to compare the students' perceptions after learning with the online corpus and online dictionary. The result of the calculation is shown in Table 5 and the significance value is .62 and p>.05. Therefore, there is no statistical difference of students' perceptions after learning using online corpus or online dictionary (U = 2007.00, p = 0,62).

Table 5

Mann-Whitney U Test of Students' Perceptions

Tests	Students' Perceptions
Mann-Whitney U	2007.00
Wilcoxon W	4152.00
Z	-.49

Asymp. Sig. (2-tailed)	.62

Discussion

As the statistical evidence confirmed, the online dictionary can have a similar effect with the online corpus when used as Data-Driven Learning (DDL) in grammar instruction due to several contributing factors. During DDL implementation with both media, every step of the DDL was effectively implemented in the online grammar classroom. The COVID-19 pandemic has turned the teaching and learning process to be fully online and it has obliged students to use mobile phones or laptops to join the lectures. It becomes an advantage in the DDL implementation because it requires access to computerized language data, in this case, access to an online corpus and an online dictionary. The researchers did not find any significant obstacles during the preparation up to the implementation of the DDL. It contradicts other studies that identified some obstacles in the preparation and implementation of DDL, as mentioned by Lin and Lee (2015), Poole (2020), and Preradović et al. (2019) that time issues on preparing DDL materials and technical challenges are found when conducting DDL activities in class.

In terms of students' participation during the teaching and learning process, the use of both media transformed the classrooms to be student-centered. At the beginning of the DDL implementation, i.e., in the searching and identification stages, the students who learned with both media actively participated in browsing the searched words and sent them to the teacher via chat in *Zoom Cloud Meetings*. Though some students mistakenly identified past perfect as present perfect at the beginning of the treatment, the teacher reminded them of the task that required them to browse sentences with the auxiliary verb *has*/*have*. Afterward, they could find new correct sentences and send them back to the teacher. The following steps of DDL were analyzing and classifying the sentences and generating rules. The students were highly motivated to analyze the parts of the sentences and come up with the rules of present perfect and passives based on their analyses and classifications. In the last step of the DDL, i.e., developing sentences, the students did the task seriously as none missed any word to develop as a sentence. Based on the researchers' observation, it could be inferred that the students' active participation in every step of the DDL generated similar effects to the use of both media. The student-centered activity of DDL that makes students active learners

is consistent with other previous studies (Lin, 2016; Lin & Lee, 2015; Smart, 2014).

Though the language data in both media are different in terms of genre, it is equally effective for performing DDL in grammar classrooms. The language data displayed in the online corpus (*SkELL*) are intended for language learning. In contrast, the online dictionary (*YourDictionary*) provides general language data deriving from various sources from fiction to non-fiction. Furthermore, the language data in both media do not contain grammatical mistakes; thus, it makes them suitable as illustrations for the students undergoing the steps of the DDL. Besides, the DDL grammar instruction requires students to identify colligation and see the word function patterns in sentence structures. Other studies have considered the use of specific and general corpora for writing instruction only (Chang, 2014; M. Chen & Flowerdew, 2018; Cotos et al., 2017); however, the current study has recognized that any genre of text that becomes language data can be used for DDL in grammar instruction.

Similarly, the different features of the online corpus and online dictionary do not contribute to the different effects of both media in grammar instruction. For example, when the students learned grammar with the online corpus (*SkELL*), they employed the *Examples* and *Word Sketch* features. The *Examples* shows the sentences containing the searched words, while the *Word Sketch* contains collocations of the searched words. On the other hand, when the students learned with the online dictionary (*YourDictionary*), they used the *Dictionary* and *Sentences* features. The *Dictionary* feature displays the meaning of the searched words and their parts of speech. The findings show that the features do not significantly contribute to the difference in students' grammar mastery. In a study conducted by Frankenberg-Garcia (2014), the definitions provided by the dictionary can expose the language learners to the target collocation or colligation; thus, the results of her study show there is no difference between those who use a dictionary to those who use corpus. In a different tone, learners who used corpus outperformed those who used dictionary in both syntactic variations and correct grammar use (Z. Huang, 2014). Though these studies do not specifically discuss the distinctive features or menus of both media, it can be inferred that the characteristics of the media can yield similar or different effects to language learning due to the uniqueness of the media.

The questionnaires show that the students have the same perceptions of both media when used in grammar instruction. Some factors that contribute to students' comparable perceptions towards both media and the discussion in the following paragraphs are based on the dimensions of the questionnaires, namely the language data, students' knowledge and ability, and their learning experience.

In terms of language data, the students in the current study approve that they can access various and authentic examples of sentences when they learn grammar using the online corpus and online dictionary. Both media have stated that the language data constructed are from large samples of texts and comprehensive reference sources. Online language data or authentic materials have been recognized by Türkmen and Aydin (2016) to help and promote the learning and teaching of grammar. The examples of sentences in *YourDictionary* are varied and numerous. This online dictionary can provide up to hundreds of sentences from various genres with different types of sentences, i.e., simple, complex, and compound. Such numerous concordance lines that can present various usages of the target words or phrases are also advantageous to DDL, as Luo (2016) advised in her study. Though *SkELL* could display only forty sentences and mainly in the form of simple sentences, the difference of sentence numbers that both websites can provide does not affect the students' perceptions. It means that the students can still use sentences from *SkELL* though the numbers are limited, and they do not hinder learning. It is also consistent with the report of Hirata Hirata (2019) that the number of sample sentences in *SkELL* is appropriate, and it could help learners understand how the target words and grammar rules are used in context. Moreover, learners may find it overwhelming to navigate and analyze a large amount of English language data (Boulton & Tyne, 2015; Moon & Oh, 2017).

Concerning the knowledge and ability that the students acquire when they learn with the online corpus and online dictionary, they find the two media useful. The students admitted that they could identify sentence patterns and practiced their analysis skills when observing the language data. Both media could provide sentences in complete that the parts of the sentences are noticeable, such as the subject, predicate, object, or complement. In the current study, these sentences are essential because the students could generate rules based on the searched words (past participles). For instance, when they learned the present perfect tenses,

the students could generate rules on which subjects would colligate with the auxiliary verb *has*/*have*, or the difference of time markers when the sentences used *since* or *for*. Sentence identification is crucial when the students learn with the DDL. It has been confirmed in previous studies that when language learners did learning tasks by discussing examples with their classmates, they made efforts to discover the patterns (Z. Huang, 2014; Smart, 2014). Furthermore, they also acknowledge that they could write better English sentences due to learning with the two media. It concurs with other studies that language learners could produce more meaningful sentences because they are exposed to the examples of language through analyzing the concordance lines (Boontam & Phoocharoensil, 2018; Z. Huang, 2014; Larsen-Walker, 2017).

Concerning the learning experience, most students positively view the significance of both media in the classrooms. The students agreed that learning with the two media was fun and motivating. During the DDL implementation, the students were actively involved in the learning process by identifying, analyzing, generating theories, and developing sentences. The exciting part of the step was when they enthusiastically volunteered to analyze the sentences. Such atmosphere has supported the findings of previous studies that language learners enjoyed making grammar rule discoveries based on the corpus data provided, as this enabled them to be proud of their learning outcomes (Geluso & Yamaguchi, 2014; Lin, 2016; Moon & Oh, 2017; Phoocharoensil, 2012). This excitement in learning may result in the students' statements that they can remember the grammar rules previously learned in class. They think that this type of task helps them better retain the acquired patterns (Akpinar et al., 2015; Alsaraireh, 2021; Gilquin, 2020; Z. Huang, 2014; Nugraha et al., 2017). Consequently, most students want to learn other grammar materials through DDL in the future as they consider this method a helpful resource for learning new grammatical knowledge (Boontam & Phoocharoensil, 2018; L. A. Huang, 2012; Lin, 2016).

Conclusion

The online corpus and online dictionary have the potential to be used as DDL in grammar instruction despite their different features and the present study has proven that both media could yield similar effects to students' grammar mastery. During the treatment with both media, every

step of DDL was effectively conducted without any significant hindrance in the online grammar classrooms, thus promoting a student-centered learning process. Although the language data in both media differs in terms of genre and the number of example sentences provided, the students can use and benefit from it during the DDL analysis stage. In addition, the distinct features of both media can aid them during the sentence production stage. When questioned about their perceptions on the use of both media in grammar learning, the students approve that they can access various and authentic examples of sentences from both media. They also find the two media useful since they may utilize the language data to discover sentence patterns and improve their analysis skills. In terms of the learning experience, the students value the importance of both media in the grammar classrooms since they believe learning with both media to be enjoyable and motivating. Statistically, the students equally perceive the use of online corpus and online dictionary in grammar learning with DDL.

Considering that the online corpus and online dictionary like *SkELL* and *YourDictionary* can be used to teach and learn grammar, it is suggested to English language teachers to use these media as DDL in grammar classrooms. During post-pandemic, teachers can still use these media to teach grammar online by combining video conferencing software and the two media. It will be efficient compared to conventional classrooms with whiteboards, colorful markers, and papers. Teachers should also familiarize themselves with the steps of DDL and the software to avoid technical problems. Furthermore, English language learners could benefit from DDL because the activity in this learning method is student-centered. Through the steps of DDL, learners practice identifying and analyzing sentence patterns and generating rules. As a result, they will hone their analytical skill, which can be applied to other grammar rules or colligations. Finally, future researchers are encouraged to replicate with a bigger sample size of EFL participants to increase the generalizability of the results. The use of other software and applications is advised to support each stage of DDL. Future researchers are also suggested to explore the long-term effects of DDL on English language learners and learners' autonomy after they learn with DDL.

Acknowledgements

This work was supported by a grant from Direktorat Riset dan Pengabdian Masyarakat (DRPM) of Badan Riset dan Inovasi National (BRIN) of Indonesia's Ministry of Education, Culture, Research, and Technology (2021).

About the Authors

Akhmad Fauzanz: A PhD student in Graduate Program of English Language Teaching, Faculty of Letters, Universitas Negeri Malang, Indonesia. He is a lecturer in Undergraduate Program of English Language Teaching, Universitas Palangka Raya, Indonesia.

Yazid Basthomi: A Professor of Applied Linguistics with the Department of English, Faculty of Letters, Universitas Negeri Malang, Indonesia. His research interests include genre analysis, corpus research, and corpus-based teaching.

Francisca Maria Ivone: An Assistant Professor in Graduate Program of English Language Teaching, Faculty of Letters, Universitas Negeri Malang, Indonesia. Her research interests include Applied Linguistics, Computer-Assisted Language Learning (CALL), Technology-Enhanced Language Learning (TELL), Autonomous Learning, Self-Directed Learning, English Language Teaching (ELT), Listening, Reading, Extensive Reading/Listening/Viewing, and Cooperative Learning.

References

Ackerley, K. (2017). Effects of corpus-based instruction on phraseology in learner English. *Language Learning and Technology*, *21*(3), 195–216. https://www.lltjournal.org/item/3006

Akpinar, K. D., Aşik, A., & Vural, A. S. (2015). The relationship between the effectiveness of vocabulary presentation modes and learners' attitudes: Corpus based contextual guessing, dictionary use and online instruction. *Asian EFL Journal*, *17*(1), 90–116. http://www.asian-efl-journal.com

Alsaraireh, M. Y. (2021). Corpus-based instruction: Fostering EFL learning in Jordan. *Asian EFL Journal*, *28*(2.3), 299–313. http://www.asian-efl-journal.com

Alzu'bi, M. A. (2015). Effectiveness of inductive and deductive methods in

teaching grammar. *Advances in Language and Literary Studies*, *6*(2), 187–193. https://doi.org/10.7575/aiac.alls.v.6n.2p.187

Aşık, A., Vural, A. Ş., & Akpınar, K. D. (2016). Lexical awareness and development through data driven learning: Attitudes and beliefs of EFL learners. *Journal of Education and Training Studies*, *4*(3), 87–96. https://doi.org/10.11114/jets.v4i3.1223

Bardovi-Harlig, K., Mossman, S., & Su, Y. (2017). The effect of corpus-based instruction on pragmatic routines. *Language Learning and Technology*, *21*(3), 76–103. https://www.lltjournal.org/item/3009

Basturkmen, H. (2018). Explicit versus implicit grammar knowledge. In J. I. Liontas (Ed.), *The TESOL Encyclopedia of English Language Teaching* (pp. 1–6). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118784235.eelt0060

Boontam, P., & Phoocharoensil, S. (2018). Effectiveness of English preposition learning through data-driven learning (DDL). *3L: Language, Linguistics, Literature*, *24*(3), 125–141. https://doi.org/10.17576/3L-2018-2403-10

Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, *67*(2), 348–393. https://doi.org/10.1111/lang.12224

Boulton, A., & Tyne, H. (2015). Corpus-based study of language and teacher education. In M. Bigelow & J. Ennser-Kananen (Eds.), *The Routledge Handbook of Educational Linguistics* (pp. 301–312). Routledge. https://doi.org/10.4324/9781315797748-29

Chambers, A. (2010). What is data-driven learning? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 345–358). Routledge.

Chang, J.-Y. (2014). The use of general and specialized corpora as reference sources for academic English writing: A case study. *ReCALL*, *26*(2), 243–259. https://doi.org/10.1017/S0958344014000056

Chen, M., & Flowerdew, J. (2018). Introducing data-driven learning to PhD students for research writing purposes: A territory-wide project in Hong Kong. *English for Specific Purposes*, *50*(April 2018), 97–112. https://doi.org/10.1016/j.esp.2017.11.004

Chen, Y. (2016). Dictionary use for collocation production and retention: A call-based study. *International Journal of Lexicography*, *30*(2), 225–251. https://doi.org/10.1093/ijl/ecw005

Chiu, L.-L., & Liu, G.-Z. (2013). Effects of printed, pocket electronic, and online dictionaries on high school students' English vocabulary retention. *Asia-Pacific Education Researcher*, *22*(4), 619–634. https://doi.org/10.1007/s40299-013-0065-1

Cotos, E., Link, S., & Huffman, S. (2017). Effects of DDL technology on genre learning. *Language Learning and Technology*, *21*(3), 104–130. https://www.lltjournal.org/item/3011

Crosthwaite, P. (2017). Retesting the limits of data-driven learning: Feedback and error correction. *Computer Assisted Language Learning*, *30*(6), 447–473. https://doi.org/10.1080/09588221.2017.1312462

Daskalovska, N. (2015). Corpus-based versus traditional learning of collocations. *Computer Assisted Language Learning*, *28*(2), 130–144. https://doi.org/10.1080/09588221.2013.803982

Dilenschneider, R. F. (2017). Examining the conditions of using an on-line dictionary to learn words and comprehend texts. *ReCALL*, *30*(1), 1–20. https://doi.org/10.1017/S0958344017000234

Dörnyei, Z., & Taguchi, T. (2010). *Questionnaires in second language research: Construction, administration, and processing.* (2nd ed.). Routledge.

Dwaik, R. A. A. (2015). English digital dictionaries as valuable blended learning tools for Palestinian college students. *English Language Teaching*, *8*(11), 1–10. https://doi.org/10.5539/elt.v8n11p1

Fleming, D. (2018). Deductive versus inductive teaching. In J. I. Liontas (Ed.), *The TESOL Encyclopedia of English Language Teaching* (pp. 1–7). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118784235.eelt0052

Flowerdew, L. (2012). *Corpora and language education*. Palgrave Macmillan.

Frankenberg-Garcia, A. (2012). Learners' use of corpus examples. *International Journal of Lexicography*, *25*(3), 273–296. https://doi.org/10.1093/ijl/ecs011

Frankenberg-Garcia, A. (2014). The use of corpus examples for language comprehension and production. *ReCALL*, *26*(2), 128–146. https://doi.org/10.1017/S0958344014000093

Friginal, E. (2013). Developing research report writing skills using corpora. *English for Specific Purposes*, *32*(4), 208–220. https://doi.org/10.1016/j.esp.2013.06.001

Gao, Z.-M. (2011). Exploring the effects and use of a Chinese-English parallel concordancer. *Computer Assisted Language Learning*, *24*(3), 255–275. https://doi.org/10.1080/09588221.2010.540469

Geluso, J., & Yamaguchi, A. (2014). Discovering formulaic language through data-driven learning: Student attitudes and efficacy. *ReCALL*, *26*(2), 225–242. https://doi.org/10.1017/S0958344014000044

Gilquin, G. (2020). Using corpora to foster L2 construction learning: A data-driven learning experiment. *International Journal of Applied Linguistics (United Kingdom)*, *February 2019*, 1–19. https://doi.org/10.1111/ijal.12317

Girgin, U. (2019). The effectiveness of using corpus-based activities on the learning of some phrasal-prepositional verbs. *TOJET: The Turkish Online Journal of Educational Technology*, *18*(1), 118–125. http://files.eric.ed.gov/ fulltext/EJ1201799.pdf

Gordani, Y. (2013). The effect of the integration of corpora in reading comprehension classrooms on English as a Foreign Language learners' vocabulary development. *Computer Assisted Language Learning*, *26*(5), 430–445. https://doi.org/10.1080/09588221.2012.685078

Hirata, Y., & Hirata, Y. (2019). Applying 'Sketch Engine for Language Learning' in the Japanese English classroom. *Journal of Computing in Higher Education*, *31*(2), 233–248. https://doi.org/10.1007/s12528-019-09208-z

Huang, L. A. (2012). The effectiveness of a corpus-based instruction in deepening EFL learners' knowledge of periphrastic causatives. *TESOL International Journal*, *6*(June), 83–108. https://www.tesol-international-journal.com/volume-6-june-2012/

Huang, Z. (2014). The effects of paper-based DDL on the acquisition of lexico-grammatical patterns in L2 writing. *ReCALL*, *26*(2), 163–183. https://doi.org/10.1017/S0958344014000020

Johns, T. (1991). Should you be persuaded: Two examples of data-driven learning. *ELR Journal*, *4*, 1–16. https://lexically.net/wordsmith/ corpus_linguistics_links/Tim Johns and DDL.pdf.

Karras, J. N. (2016). The effects of data-driven learning upon vocabulary acquisition for secondary international school students in Vietnam. *ReCALL*, *28*(2), 166–186. https://doi.org/10.1017/S0958344015000154

Lai, S.-L., & Chen, H.-J. H. (2015). Dictionaries vs concordancers: Actual practice of the two different tools in EFL writing. *Computer Assisted Language Learning*, *28*(4), 341–363. https://doi.org/10.1080/09588221.2013.839567

Laosrirattanachai, P., & Ruangjaroon, S. (2021). Corpus-based creation of tourism, hotel, and airline business word lists. *LEARN Journal: Language Education and Acquisition Research Network*, *14*(1), 50–86.

Larsen-Walker, M. (2017). Can Data Driven Learning address L2 writers' habitual errors with English linking adverbials? *System*, *69*(October 2017), 26–37. https://doi.org/10.1016/j.system.2017.08.005

Lee, H., Warschauer, M., & Lee, J. H. (2019). The effects of corpus use on second language vocabulary learning: A multilevel meta-analysis. *Applied Linguistics*, *40*(5), 721–753. https://doi.org/10.1093/applin/amy012

Li, S. (2017). Using corpora to develop learners' collocational competence. *Language Learning and Technology*, *21*(3), 153–171. https://www.lltjournal.org/item/3018

Lin, M. H. (2016). Effects of corpus-aided language learning in the EFL grammar classroom: A case study of students' learning attitudes and teachers' perceptions in Taiwan. *TESOL Quarterly*, *50*(4), 871–893. https://doi.org/10.1002/tesq.250

Lin, M. H., & Lee, J.-Y. (2015). Data-driven learning: Changing the teaching of grammar in EFL classes. *ELT Journal*, *69*(3), 264–274. https://doi.org/10.1093/elt/ccv010

Luo, Q. (2016). The effects of data-driven learning activities on EFL learners' writing development. *SpringerPlus*, *5*(1255), 1–13. https://doi.org/10.1186/s40064-016-2935-5

Mizumoto, A., & Chujo, K. (2016). Who is data-driven learning for? Challenging the monolithic view of its relationship with learning styles. *System*, *61*(2016), 55–64. https://doi.org/10.1016/j.system.2016.07.010

Mizumoto, A., Chujo, K., & Yokota, K. (2016). Development of a scale to measure learners' perceived preferences and benefits of data-driven learning. *ReCALL*, *28*(2), 227–246. https://doi.org/10.1017/S0958344015000208

Moon, S., & Oh, S.-Y. (2017). Unlearning overgenerated be through data-driven learning in the secondary EFL classroom. *ReCALL*, *30*(1), 48–

67. https://doi.org/10.1017/S0958344017000246

Mueller, C. M., & Jacobsen, N. D. (2016). A comparison of the effectiveness of EFL students' use of dictionaries and an online corpus for the enhancement of revision skills. *ReCALL*, *28*(1), 3–21. https://doi.org/10.1017/S0958344015000142

Nugraha, S. I., Miftakh, F., & Wachyudi, K. (2017). *Teaching grammar through Data-Driven Learning (DDL) approach*. *82*(Conaplin 9), 300–303. https://doi.org/10.2991/conaplin-16.2017.68

Nunan, D. (1998). Teaching grammar in context. *ELT Journal*, *52*(2), 101–109. https://doi.org/10.1093/elt/52.2.101

Nurmukhamedov, U. (2012). Online English-English learner dictionaries boost word learning. *English Teaching Forum*, *50*(4), 10–15. https://americanenglish.state.gov/files/ae/resource_files/50_4_4_nurmukhamedov.pdf.

Pawlak, M. (2018). Grammar learning strategy inventory (GLSI): Another look. *Studies in Second Language Learning and Teaching*, *8*(2), 351–379. https://doi.org/10.14746/ssllt.2018.8.2.8

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

Poole, R. (2016). A corpus-aided approach for the teaching and learning of rhetoric in an undergraduate composition course for L2 writers. *Journal of English for Academic Purposes*, *21*(March 2016), 99–109. https://doi.org/10.1016/j.jeap.2015.12.003

Poole, R. (2020). "Corpus can be tricky": Revisiting teacher attitudes towards corpus-aided language learning and teaching. *Computer Assisted Language Learning*, *0*(0), 1–22. https://doi.org/10.1080/09588221.2020.1825095

Preradović, N. M., Posavec, K., & Unić, D. (2019). Corpus-supported foreign language teaching of less commonly taught languages. *International Journal of Instruction*, *12*(4), 335–352. https://doi.org/10.29333/iji.2019.12422a

Quinn, C. (2014). Training L2 writers to reference corpora as a self-correction tool. *ELT Journal*, *69*(2), 165–177. https://doi.org/10.1093/elt/ccu062

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

https://doi.org/10.1017/S0958344014000081

Sripicharn, P. (2010). How can we prepare learners for using language corpora? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 371–384). Routledge.

Tananuraksakul, N. (2015). The effect of online dictionaries usage on EFL undergraduate students' autonomy. *Teaching English with Technology*, *15*(4), 3–15.

Tono, Y., Satake, Y., & Miura, A. (2014). The effects of using corpora on revision tasks in L2 writing with coded error feedback. *ReCALL*, *26*(2), 147–162. https://doi.org/10.1017/S095834401400007X

Tsai, K.-J. (2019). Corpora and dictionaries as learning aids: inductive versus deductive approaches to constructing vocabulary knowledge. *Computer Assisted Language Learning*, *32*(8), 805–826. https://doi.org/10.1080/09588221.2018.1527366

Türkmen, Y., & Aydin, S. (2016). The effects of using online concordancers on teaching grammar. *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi Nisan*, *20*(1), 145–152. http://files.eric.ed.gov/fulltext/ ED565621.pdf

Yoon, C. (2016). Concordancers and dictionaries as problem-solving tools for ESL academic writing. *Language Learning and Technology*, *20*(1), 209–229. https://doi.org/10125/44453