doi：10. 20776/S13482084-70-P289

千葉大学教育学部研究紀要　第70巻　289～297頁（2022）

Development of an Online DDL Tool for Secondary School Learners

NISHIGAKI Chikako[1)]*, AKASEGAWA Shiro[2)] and OGHIGIAN Kathryn[3)]

[1)]Faculty of Education, Chiba University, Japan
[2)]Lago Institute of Language, Japan
[3)]Vantage College, University of British Columbia, Canada

中学生・高校生のためのオンライン・データ駆動型学習ツールの開発

西垣知佳子[1)]*・赤瀬川史朗[2)]・オヒガン・キャサリン[3)]

[1)]千葉大学・教育学部
[2)]Lago言語研究所
[3)]ブリティッシュコロンビア大学・バンテージ

This paper reports on the application of data-driven learning（DDL）in secondary schools in Japan. DDL has been used primarily with higher proficiency level post-secondary school learners; the goal of this research is to determine how the benefits of DDL can be achieved with younger, less proficient learners. First, we report on a DDL tool we developed called hDDL. We explain how it was developed and how it works. Second, we verify the level-appropriateness of the hDDL pedagogical corpus. Third, we report on feedback we collected from secondary school learners after using hDDL to learn a specific grammar task. We found that hDDL contains a level-appropriate corpus for Japanese secondary school learners based on Japanese Ministry of Education guidelines and CERF-J parameters. Learner feedback suggested that hDDL can be useful for inductive learning with this target population. We hope to continue to expand its features and offer it for use with international learners.

データ駆動型学習（data-driven learning：DDL）とは，コーパスと検索ツールを使って行う外国語の学習方法の１つである。DDLは海外で活用が広がっているものの，その対象は，外国語の習熟度が中級・上級の大学生学習者である。本研究者グループでは，登録不要，無料，著作件フルーで使える入門期・初級レベルのDDLツールを日本の中学生，高校生用に開発し，公開した。本稿では，はじめにhDDLの開発プロセスと学習機能を紹介する。続いて，hDDLが搭載する教育用コーパスのレベルを検証し，さらに中学生がDDLツールを使用して学習した後に集めたフィードバックを報告する。結果として，hDDLは学習指導要領とCERF-Jのパラメータに基づいて，日本の中学・高校の学習者のレベルに適切であること，また，学習者のフィードバックから，hDDLは帰納的学習に有用であること等が示唆された。今後はhDDLが搭載するコーパスと機能を拡充し，さらには英語だけでツールを利用できるように英語版を作成し，海外の学習者にも利用できるようにしていく。

キーワード：data-driven learning（データ駆動型学習），pedagogical corpus（教育用コーパス），
introductory-level（入門期レベル），noticing（発見活動）

1. Introduction

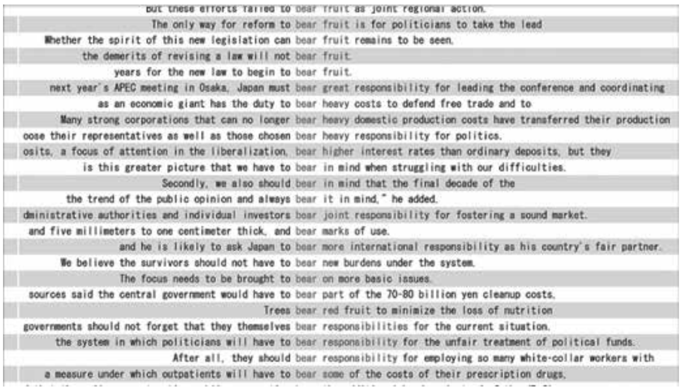
1.1 A Corpus-Based Method for Learning and Teaching English

With the development of corpus linguistics across the last three decades, there have been a growing number of uses for corpora（electronic language databases）in language research and language teaching. Corpora have been used by researchers and educators as language data to compile dictionaries, create educational vocabulary lists, and produce language textbooks, among other applications. Johns（1991）proposed data-driven learning（DDL）in which a corpus itself is searched by learners and the search results become the basis for language study material. With this approach, learners are exposed to target language directly by noticing language patterns in numerous examples in a corpus. The former is referred to as an indirect use of corpus, and the latter, as a direct use of corpus（Leech, 1997; Römer, 2006）.

With DDL, the direct use of a corpus typically has learners use search software to examine concordance lines（partial sentences）within a corpus that contain the target word(s). Figure 1 shows an example of a search result using an English newspaper corpus. In this case, the target word is *bear,* which is a word Jap-

*連絡先著者：西垣知佳子　gaki@faculty.chiba-u.jp

anese English learners are familiar with from their English textbooks. The keyword, *bear,* can be seen displayed in the center of the screen vertically in its original contexts. This format of concordance lines is called KWIC (Keyword in Context). Typically, DDL software also includes a sorting function which can be used to highlight one or more words to the right or left of the KWIC. This makes patterns easier to notice, which is the first step in deducing language rules. In Figure 1, learners can see that *bear* is often used as a verb rather than a noun, as in *bear fruit, bear responsibility,* and *bear in mind*. The use of DDL encourages learners to actively observe and analyze their search results and learn lexical and grammatical rules by discovering and exploring inductively and inquisitively. Johns (1991) writes that the "language-learner is also, essentially, a research worker whose learning needs to be driven by access to linguistic data–hence the term 'data-driven learning' (DDL) to describe the approach" (p. 2).



```
             But these efforts failed to bear fruit as joint regional action.
            The only way for reform to bear fruit is for politicians to take the lead
Whether the spirit of this new legislation can bear fruit remains to be seen.
       the demerits of revising a law will not bear fruit.
          years for the new law to begin to bear fruit.
  next year's APEC meeting in Osaka, Japan must bear great responsibility for leading the conference and coordinating
       as an economic giant has the duty to bear heavy costs to defend free trade and to
  Many strong corporations that can no longer bear heavy domestic production costs have transferred their production
oose their representatives as well as those chosen bear heavy responsibility for politics.
osits, a focus of attention in the liberalization, bear higher interest rates than ordinary deposits, but they
      is this greater picture that we have to bear in mind when struggling with our difficulties.
               Secondly, we also should bear in mind that the final decade of the
     the trend of the public opinion and always bear it in mind," he added.
dministrative authorities and individual investors bear joint responsibility for fostering a sound market.
 and five millimeters to one centimeter thick, and bear marks of use.
           and he is likely to ask Japan to bear more international responsibility as his country's fair partner.
  We believe the survivors should not have to bear new burdens under the system.
          The focus needs to be brought to bear on more basic issues.
 sources said the central government would have to bear part of the 70-80 billion yen cleanup costs.
                                  Trees bear red fruit to minimize the loss of nutrition
governments should not forget that they themselves bear responsibilities for the current situation.
   the system in which politicians will have to bear responsibility for the unfair treatment of political funds.
                 After all, they should bear responsibility for employing so many white-collar workers with
  a measure under which outpatients will have to bear some of the costs of their prescription drugs.
```

Figure 1　*Screenshot of a Search Result for* bear

In addition, this direct use of DDL has the advantage of exposing students to a large amount of authentic English, and this type of inductive learning has been shown to be very effective (see Boulton & Cobb, 2017). Aspects of DDL are consistent with many of the major theories of second language acquisition, such as consciousness-raising, noticing, salience, discovery learning, meta-cognitive skills, autonomy, individualization, learner-centeredness, critical thinking, focus on form, task-based learning, usage-based learning, and constructivism, among others.

The corpora traditionally used as DDL learning materials are primarily collections of authentic language used by adult native speakers, the most common of which include the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA). Due to the language level of these corpora, Johns (1991), the original proponent of DDL, judged DDL to be a learning method restricted to intermediate and advanced learners. More recently, although Wicher (2020) has recognized that "[d]ata-driven learning (DDL) represents one of the most promising applications" of corpora (p. 31), he also points out that its success has been limited to university learners with advanced foreign language proficiency, and that the secondary school classroom remains comparatively uncharted territory. Unsurprisingly, it is uncommon for young school-age learners to use corpora in second language acquisition during their primary or secondary school education. In addition to the high-level language and complex grammar found in corpora, Pérez-Paredes (2020) suggests teachers' lack of familiarity with corpus resources and a lack of appropriate resources for pre-tertiary groups of learners as additional reasons. In their comprehensive meta-analysis of corpus studies, Boulton and Cobb (2017) cited only 10 of the 88 cases they reviewed as having been designed for use at the secondary school level. The use of corpora for lower proficiency learners has not developed in recent years at the same pace as that of less specialized applications for corpora.

Two key measures are needed to solve the problems that prohibit a wider use of DDL at the secondary school level. The first measure is the creation of a *pedagogical* corpus that is level-appropriate for beginner learners. (For more on pedagogical corpora, see Flowerdew, 2009 and Timmis, 2015.) The second measure is the development of a user-friendly DDL tool that is simple to use for both secondary school learners and teachers.

1.2　Objectives

The purpose of this paper is (a) to report on the development of a novel pedagogical corpus and DDL tool targeted to secondary school learners and teachers who have no special knowledge of corpora and provide an evaluation of its appropriateness based on existing Ministry of Education guidelines and the Common European Framework of Reference for Languages developed for Japanese learners (CEFR-J); and (b) to implement this corpus and tool (called hDDL) with Japanese junior high school students to elicit their feedback on its usefulness and ease of use.

2. Development of a Pedagogical Corpus and DDL Tool for Secondary School Learners

2.1 Overview

One exception to the lack of resources for beginner-level students is the Sentence Corpus of Remedial English (SCoRE) (https://www.score-corpus.org/). SCoRE (Chujo et al., 2015) was created for low proficiency level university students. It consists of complete sentences rather than the concordance lines found in traditional corpora, and can be searched by grammatical category or lexical target word. Although SCoRE was created for university-level students, its rationale and procedures were used in this project as a basis for creating a similar pedagogical corpus and tool for secondary school learners called "hDDL" (https://h.ddl-study.org/). In this section, we outline how we developed this pedagogical corpus and tool.

2.2 Development of a Pedagogical Sentence Corpus with L1 Translation

In order to solve the first problem of expanding the use of DDL among young learners, we needed to create a pedagogical corpus for secondary school learners. To achieve this, we first created a source corpus as shown in Figure 2 as ①. The collected data was from (a) selected school textbooks published in Japan, China, Korea and Taiwan, as well U.S. reading textbooks; (b) graded readers such as the *Oxford Reading Tree*; and (c) EFL material that was available on the Internet such as *Science News for Students,* and *News in Levels*. Table 1 shows the number of tokens (words) and sentences taken from each category. The resulting source corpus included approximately 24 million words.

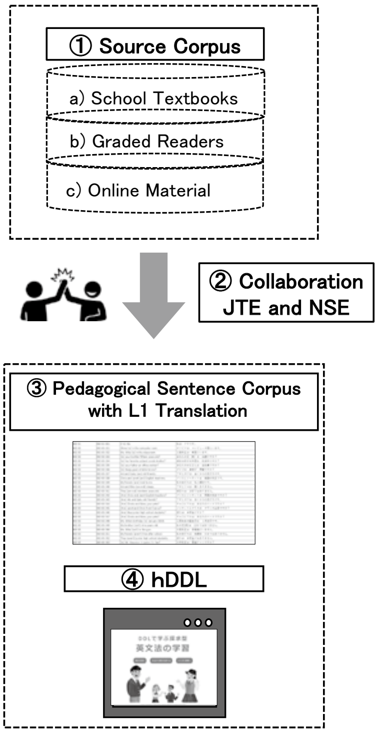

① Source Corpus
a) School Textbooks
b) Graded Readers
c) Online Material
② Collaboration JTE and NSE
③ Pedagogical Sentence Corpus with L1 Translation
④ hDDL

Figure 2 *Process of the Development DDL Tool for Secondary School Learners*

Table 1 *Source Corpus*

Types of Language Data	*N* of Tokens (Words)	*N* of Sentences (Sentences)
a) school textbooks	736,703	88,404
b) graded readers	300,339	41,753
c) online resources	22,990,173	2,009,209
Total	24,027,215	2,139,366

After we compiled the 24-million-word source corpus, Japanese teachers of English (JTE) and native speakers of English (NSE) (shown in Figure 2 as ②) collaborated to create a pedagogical sentence corpus (shown as ③), for hDDL (④). Using a corpus search tool[1], the teachers extracted and selected sentences from the source corpus, and using these as examples, rewrote or modified them with simpler vocabulary and in relatable contexts while retaining the original grammatical structure. Because one key factor that impacts sentence comprehension is sentence length (Yano et al., 1994), there was an emphasis placed on creating short and simple sentences, and this was given special consideration.

Many of the short sentences found in the source corpus were appropriate as they were (e.g., *I like baseball. I have a sister.*); however, some sentences contained culturally-related names or themes unknown to the target learners (e.g., *I like baclava.*). Expressions and unfamiliar words were excluded or modified. In addition, age-related interests were considered. For example, to modernize the hDDL content, sentences referring to modern technology such as mobile phones, websites, SNS, and video games were included. We also developed illustrated characters (secondary school students, their parents, siblings, school teachers, friends, and pets) to create a background and storyline for the English sentences. In this way, the English sentences in hDDL have a certain degree of thematic coherence. The end result was the creation of a collection of level-appropriate, original, and copyright-free sentences.

Finally, another unique feature hDDL provides is a parallel L1 translation for each sentence used in the L2 corpus (shown as ③ in Figure 2). When a sentence contains an unknown word, learners can easily check the meaning by reviewing its L1 translation. This lessens the learning burden so learners can focus on the discovery of language patterns. In this way, the pedagogical sentence corpus provides an intrinsic and critical scaffolding benefit.

The pedagogical sentences for hDDL were developed and tagged according to grammar items such as passives and participles. Table 2 provides a list of 17 grammar items currently (as of October, 2021) included in hDDL.

Table 2　*17 Grammar Items in hDDL*

1 be-verbs	7 auxiliaries	13 comparisons
2 progressives	8 subjunctive	14 relative pronouns
3 past	9 sentence patterns	15 relative adverbs
4 future	10 infinitives	16 conjunctions
5 perfect forms	11 gerunds	17 indirect questions
6 passives	12 participles	

The hDDL pedagogical corpus currently consists of 13,756 words (when contractions such as *I'm* are counted as one word), or 14,609 words (if counted as two), and 2,361 sentences.

2.3　Development of the DDL Tool

This web-based DDL tool can be accessed without cost or registration. There are three subtools within the hDDL which are similar to SCoRE, and these are explained in this section.

2.3.1　hDDL Homepage

Figure 3 shows the hDDL homepage. The English translation of the title, shown as ❶ is "Data-Driven Learning: Explore English Grammar." The illustrations on the page are cut-out papercraft originally produced for hDDL. The two main characters are a boy and girl. By clicking the appropriate button in the middle of the page, learners can select one of the three subtools they want to use: Pattern Search (❷), Word Search (❸), or Try the Quiz (❹). Pattern Search is a pattern browser which allows learners to choose a grammar item to focus on. Word Search is a concordancer that displays search results of a target word or words. Try the Quiz is a word arrangement quiz; it automatically makes and marks questions. A game-like quiz was included to motivate learners and identify learners' weaknesses in grammar.

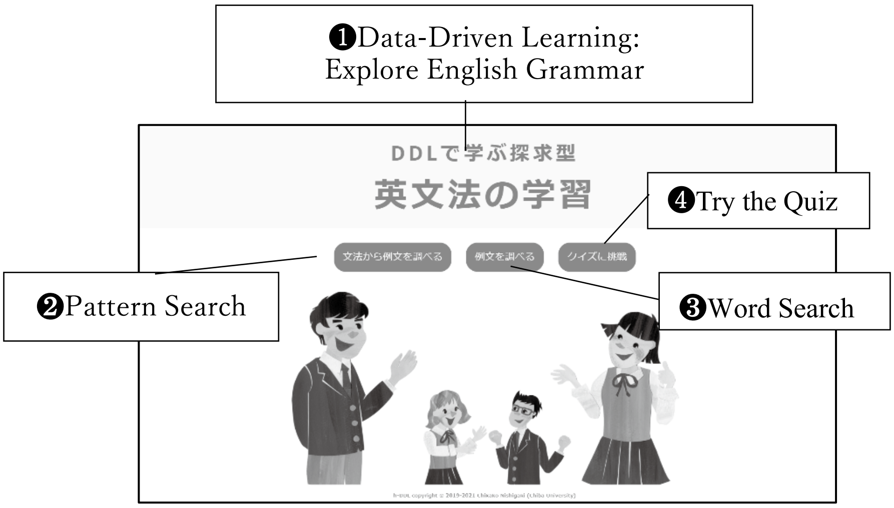


Figure 3　*Screenshot of the hDDL Homepage*

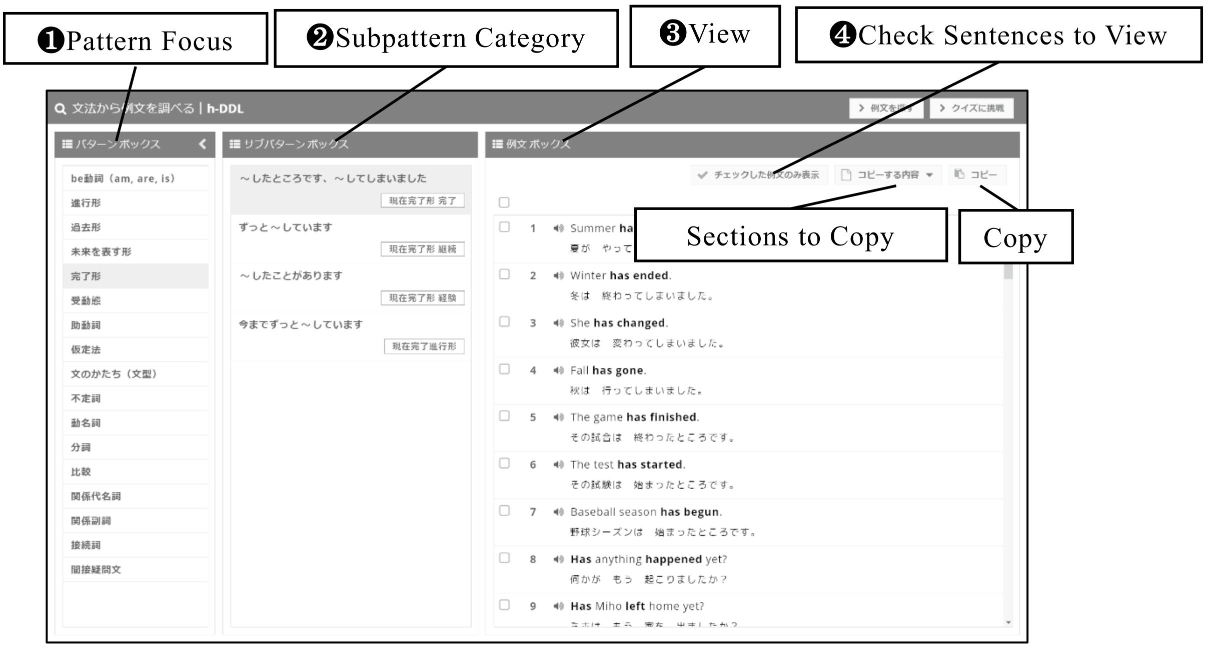
2.3.2　Pattern Search Subtool

Figure 4 shows the Pattern Search subtool screen. As can be seen, the layout of the Pattern Search subtool screen consists of three vertical spaces. The leftmost space, Pattern Focus, is for the learner to select the grammar item they want to study (❶). The second space from the left displays subcategories of grammatical items (❷). For example, if learners select the perfect tense, there are four subcategories: perfect, continuous, experience, and present perfect progressive. Search results are shown under View (❸).

2.3.3　Word Search Subtool

Figure 5 shows a screenshot of the Word Search page. On the left side of the page (❶Search), options are provided for search conditions. Search results appear in the area to the right, under View (❷). A learner can type a search term or terms (e.g., *like* or *like to*) in the Search Box (❸), and can choose the length of the sentence they are searching for (❹). If they select a shorter sentence, the structure of the searched sentences is likely to be simpler. For example, if a learner chooses sentences with three words, they will get *I ate pizza,* and if they pick a five-word sentence, the resulting sentences will be grammatically denser, such as *I ate pizza last night* or *My big sister ate pizza.*

One advantage of DDL is that it enables students to observe many examples at once. At the same time, multiple examples can be overwhelming and intimidat-



be動詞（am, are, is）
進行形
過去形
未来を表す形
完了形
受動態
助動詞
仮定法
文のかたち（文型）
不定詞
動名詞
分詞
比較
関係代名詞
関係副詞
接続詞
間接疑問文

～したところです、～してしまいました　現在完了形 完了
ずっと～しています　現在完了形 継続
～したことがあります　現在完了形 経験
今までずっと～しています　現在完了進行形

□
□ 1 Summer **ha**
夏が　やって
□ 2 Winter **has ended**.
冬は　終わってしまいました。
□ 3 She **has changed**.
彼女は　変わってしまいました。
□ 4 Fall **has gone**.
秋は　行ってしまいました。
□ 5 The game **has finished**.
その試合は　終わったところです。
□ 6 The test **has started**.
その試験は　始まったところです。
□ 7 Baseball season **has begun**.
野球シーズンは　始まったところです。
□ 8 **Has** anything **happened** yet?
何かが　もう　起こりましたか？
□ 9 **Has** Miho **left** home yet?

Figure 4　*Screenshot of a Pattern Search*

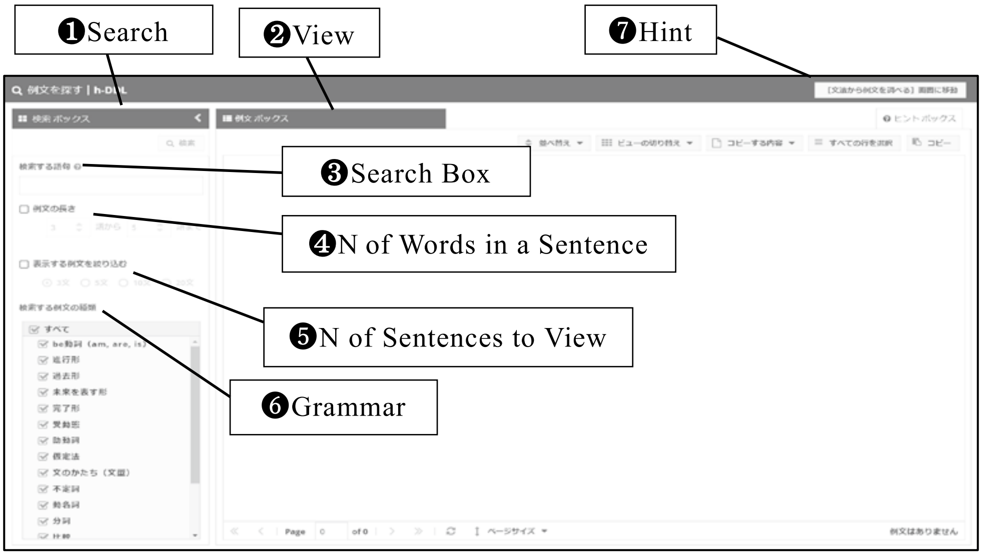
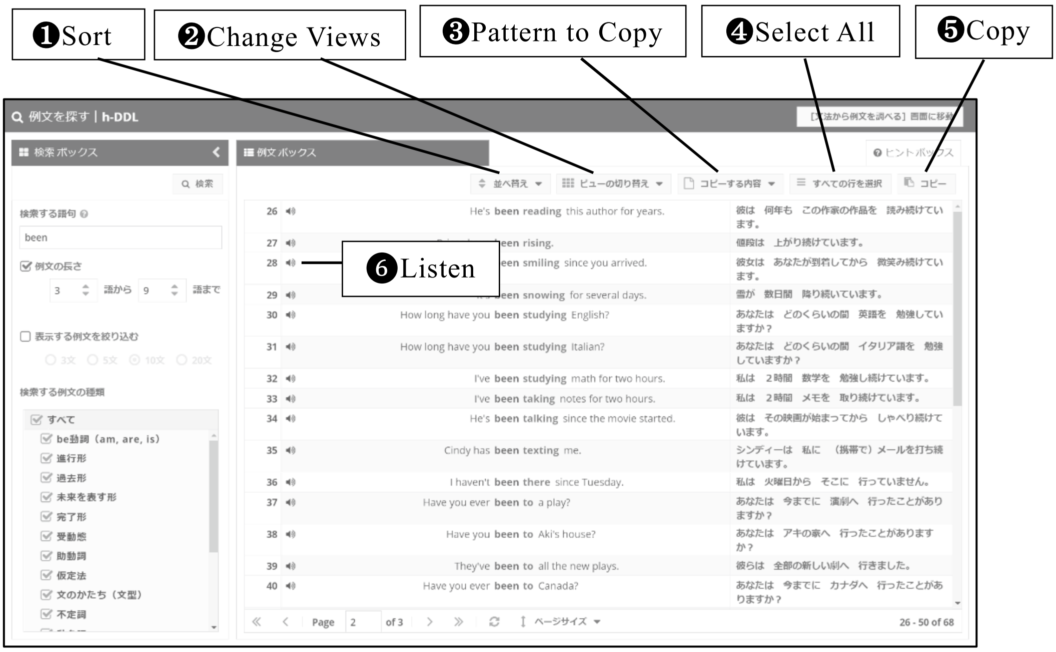


Figure 5　*Word Search*（*1*）

ing. To address this, learners can limit the number of sentences in the results on the screen to 3, 5, 10, or 20（❺). Additionally, a particular grammar focus can be selected using the Grammar button（❻); for example, if a learner enters *I'm* in the search box, *I'm* in various grammatically structured sentences will appear in the view results. The learner then has the option to further narrow the search parameters to the present progressive, which they can choose from a list of grammatical categories, to view, for example, *I'm reading*. Finally, the Hint tab（❼）guides students on how to write a search word or words and explains how to filter search results.

Figure 6 shows a screenshot of sample search results. Each separate sentence displayed in the search space is called a concordance line. Learners can listen to the pronunciation of the sentences by clicking the speaker icon to the left of each concordance line. They can also click on the Sort button（❶）to start the sorting process. Learners can sort words to the right or the left of the target word. Using the Change Views button（❷), the learner can switch between the normal sentence view and the KWIC view. The three tabs from the third left to the far right are to copy the concordance lines. The third button from the right, Pattern to Copy（❸), selects the concordance lines to copy. The second tab from the right, Select All（❹), selects all the concordance lines on the screen. The Copy button on the far right（❺）will copy the selected concordance lines so that they can be pasted into a Word or Excel file. Learners can also hear a sentence by clicking the speaker icon（❻).



例文を探す｜h-DDL　　[文法から例文を調べる] 画面に移動

検索ボックス

検索

検索する語句
been

☑ 例文の長さ
3 語から 9 語まで

□ 表示する例文を絞り込む
○ 3文 ○ 5文 ◉ 10文 ○ 20文

検索する例文の種類
☑ すべて
☑ be動詞（am, are, is）
☑ 進行形
☑ 過去形
☑ 未来を表す形
☑ 完了形
☑ 受動態
☑ 助動詞
☑ 仮定法
☑ 文のかたち（文型）
☑ 不定詞

例文ボックス　　ヒントボックス

並べ替え　ビューの切り替え　コピーする内容　すべての行を選択　コピー

26	He's **been reading** this author for years.	彼は　何年も　この作家の作品を　読み続けています。
27	**een rising.**	値段は　上がり続けています。
28	**een smiling** since you arrived.	彼女は　あなたが到着してから　微笑み続けています。
29	**een snowing** for several days.	雪が　数日間　降り続いています。
30	How long have you **been studying** English?	あなたは　どのくらいの間　英語を　勉強していますか？
31	How long have you **been studying** Italian?	あなたは　どのくらいの間　イタリア語を　勉強していますか？
32	I've **been studying** math for two hours.	私は　２時間　数学を　勉強し続けています。
33	I've **been taking** notes for two hours.	私は　２時間　メモを　取り続けています。
34	He's **been talking** since the movie started.	彼は　その映画が始まってから　しゃべり続けています。
35	Cindy has **been texting** me.	シンディーは　私に　（携帯で）メールを打ち続けています。
36	I haven't **been there** since Tuesday.	私は　火曜日から　そこに　行っていません。
37	Have you ever **been to** a play?	あなたは　今までに　演劇へ　行ったことがありますか？
38	Have you **been to** Aki's house?	あなたは　アキの家へ　行ったことがありますか？
39	They've **been to** all the new plays.	彼らは　全部の新しい劇へ　行きました。
40	Have you ever **been to** Canada?	あなたは　今までに　カナダへ　行ったことがありますか？

Page 2 of 3　ページサイズ　26 - 50 of 68

Figure 6　*Word Search*（2）

2.3.4　Try the Quiz

In *Try the Quiz,* learners click on the grammar item they want to study, and the quiz begins. They look at the L1（Japanese）and arrange the L2 words at the bottom of the screen to match the L1（see Figure 7）. When learners hover the mouse over the English word, they are quickly directed to the answer box. When they click the answer button, the correct sentence pops up. They can also hear a word by clicking the speaker icon. Questions are batched in sets of five（see Figure 8）. After completing five questions, learners are told how many correct answers they have achieved. If there are any errors, learners can jump to the *Pattern Search* page, so they can review the specific grammatical item featured in that pattern.

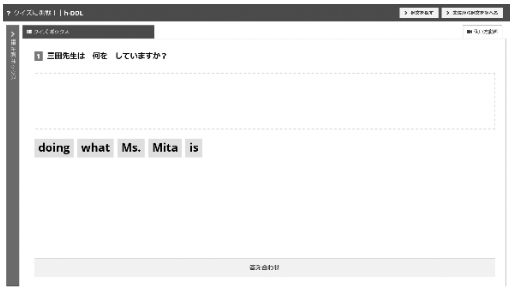
2.3.5　Search Engine

To search the hDDL corpus efficiently, we used the search engine Blacklab Query Tool（http://inl.github.io/BlackLab/）. The Blacklab Query Tool is a specialized search engine for corpus searches and uses a standard query language called Corpus Query Language（CQL）. Because it may be difficult for junior and senior high school learners to use CQL search formulas, wildcards are used instead; that is, the hDDL automatically converts the complex standard query language to simple search expressions or wildcards.

3．Evaluation of Level-Appropriateness

3.1　Method and Results

We examined the proficiency level and appropriateness of the corpus as study material in terms of（a）grammar items,（b）sentence length, and（c）vocabulary.

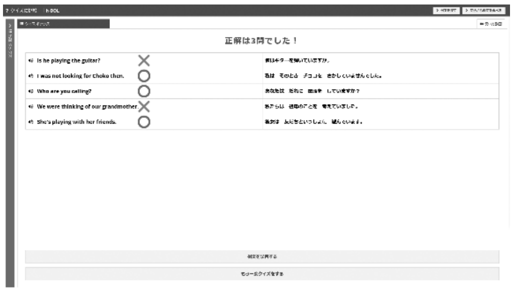


1　三田先生は　何を　していますか？

doing　what　Ms.　Mita　is

答え合わせ

Figure 7　*Screenshot* of *Try the Quiz*



正解は3問でした！

Is he playing the guitar?	✗	彼はギターを弾いていますか。
I was not looking for Choko then.	○	私は　そのとき　チョコを　さがしていませんでした。
Who are you calling?	○	あなたは　だれに　電話を　していますか？
We were thinking of our grandmother.	✗	私たちは　祖母のことを　考えていました。
She's playing with her friends.	○	彼女は　友だちといっしょに　遊んでいます。

[illegible]

もう一度クイズをする

Figure 8　*Screenshot of Try the Quiz Answer*

3.1.1　Grammar Items

The educational guidelines published by the Japanese Ministry of Education list 11 grammatical items that should be taught to Japanese junior high school learners（7th–9th graders）. These are pronouns, conjunctions, auxiliaries, prepositions, tenses and aspect, comparatives, to-infinitives, gerunds, participles, passive voice, and subjunctives. The guidelines also list eight grammatical items for Japanese high school learners（10th–12th graders）that should be taught; these are infinitives, relative pronouns, relative adverbs, conjunctions, auxiliaries, prepositions, tenses and aspects of verbs, and subjunctives. Of these 19 items, hDDL includes 17; only pronouns and prepositions are not included. Since hDDL covers nearly all of the grammatical items assigned to teach in the government guidelines, it was deemed appropriate for use at the secondary school level (junior and senior high school).

3.1.2 Sentence Length

Figure 9 shows an overview of the sentence length of the hDDL pedagogical corpus. In Figure 9, the number of words in a sentence is on the horizontal axis, and the number of sentences containing that number of words is on the vertical axis. The shortest sentence is a two-word sentence. The longest sentence has 15 words. The average sentence length is 5.82 words. There are no one-word sentences because the current version of hDDL does not include imperatives as a target grammar item. The most widely used 7th grade English textbook in Japan has a range of 3.50 words in Unit 1 to 6.16 words in the last unit. Since hDDL includes a lot of 2- and 3-word sentences and the average sentence length is less than 6.16 words, we can conclude that hDDL is appropriate as a DDL tool for introductory and beginner-level learners. Using short, simple sentences allows learners focus on discovering the targeted grammar.

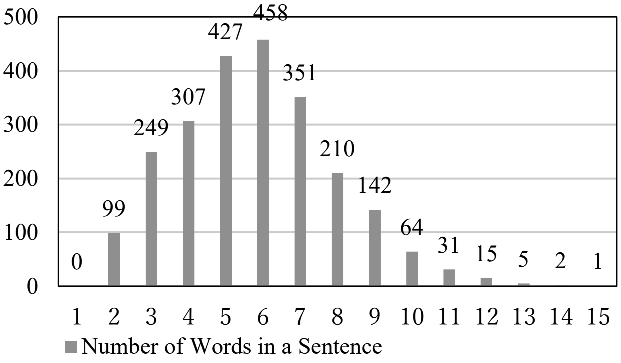


Figure 9 ***Sentence Length of hDDL***

3.1.3 Vocabulary Level

We investigated the lexical level of the hDDL pedagogical corpus using the CEFR-J. The CEFR (Common European Framework of Reference for Languages) is well-known among language teachers and is used to describe the achievements of language learners worldwide. The CEFR-J was developed specifically for Japanese learners. The CEFR uses six stages to describe learners' levels: A1, A2, B1, B2, C1, and C2 from low to high proficiency. The CEFR-J consists of an original scale of sublevels for stages A1 to B2 to more discretely categorize lower-proficiency level learners. The developers of CEFR-J had found that roughly 80% of Japanese learners are at the A level (A1 and A2), 20% are at B level (B1 and B2), and almost none are at C level. (See https://tufs-sgu.com/language_ability/cefr-j/ for the details of CEFR-J.) According to the Japanese Ministry of Education, the goal of 9th graders is A1 and 12th graders, A2.

The New Word Level Checker (NWLC) was used to evaluate the proficiency level of the hDDL vocabulary within the CERF-J framework. The NWLC is based on the CEFR-J criteria and was developed by Mizumoto (2021) (https://nwlc.pythonanywhere.com/). The results can be seen in Figure 10. Only 6.20% of the hDDL words are grouped as unknown (*bathtub*, *cellphone*, *compass*, *differ*, *doorbell*, *fingerprint*, *furry*, *grocery*, *itchy*, *license*, *mammal*, *monorail*, and others). Approximately 85% of the vocabulary in hDDL belongs to CEFR-J A1 and 90% to A2. These results suggest that hDDL is appropriate for the level of our target learners.

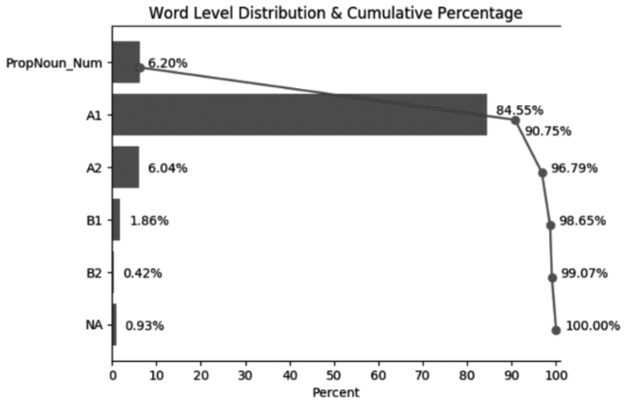


Figure 10 ***Vocabulary Level of hDDL Pedagogical Corpus***

4. Learner Feedback

hDDL was introduced to secondary school learners in previous studies. One study looked at learners' improvement in the acquisition of grammatical knowledge using hDDL. In that study, 151 8th graders examined the usage of *have to* and *has to*. They took a pre-test, post-test and a retention test. The three tests were identical. The scores statistically improved between the pre-test and post-test; retention was demonstrated four weeks after instruction (Nishigaki, Kawana, Yamaguchi et al., 2021). As a follow-up, this study focused on what learners thought of the usefulness of hDDL.

4.1 Participants

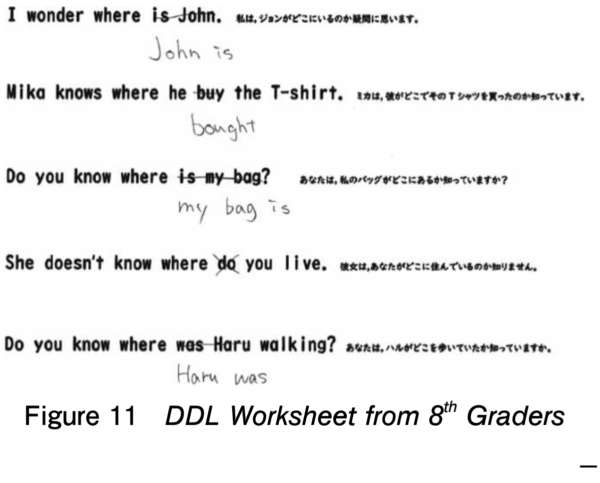
Participants were 139 first-year junior high school students (7th grade), and 70 second-year junior high school learners (8th grade) in Japan. They attend a 50-minute English class four times a week. They started learning English grammar in a communicative way from the 7th grade. Most of the students learned English through listening and speaking at elementary school from the 5th grade and some started from 1st grade. The 7th graders in this study had attended one of the previous studies (Nishigaki, Akasegawa, Kawana et al, 2021). In that study, we reported on the score increase of grammar test scores between the pre-test and the post-test. We also found that different instructions resulted in different student findings. However, in this study we focused on the learners' comments and opinions.

4.2　Target Grammar

The aim of the hDDL lessons was to target a structure of English sentences that students had not previously been taught. The target grammatical pattern was [subject + show + person + thing/things] for 7th graders and the indirect question form [subject + verb + wh-interrogative + subject + verb] for 8th graders. Students were introduced to these sentence structures for the first time in school.

4.3　Procedure

In the hDDL class, they were given a worksheet with sentences that contained the target grammatical items with errors (see Figure 11). Working individually, learners searched for the target grammatical item and compared the sentence on the worksheet with the hDDL search results. If there was an error on the worksheet, they corrected it. They also summarized the grammatical rules they found. Many of the 8th grade students were able to identify a different word order of the wh-questions (*I wonder where John is*) from sentences without the wh- question (*Where is John?*). An example is illustrated by a student in Figure 12. Next, students compared and discussed their answers on the error corrections with a partner. After the pair work, the teacher and learners reviewed the answers together. The teacher also elicited opinions on the target grammar.



I wonder where ~~is John~~. 私は，ジョンがどこにいるのか疑問に思います。

John is

Mika knows where he ~~buy~~ the T-shirt. ミカは，彼がどこでそのＴシャツを買ったのか知っています。

bought

Do you know where ~~is my bag~~? あなたは，私のバッグがどこにあるか知っていますか？

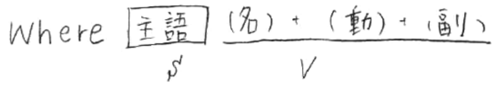
my bag is

She doesn't know where ~~do~~ you live. 彼女は，あなたがどこに住んでいるのか知りません。

Do you know where ~~was Haru~~ walking? あなたは，ハルがどこを歩いていたか知っていますか。

Haru was

Figure 11　*DDL Worksheet from 8th Graders*



Where 主語 (名) + (動) + (副)

S V

Figure 12　*Learner's Note from Worksheet*

4.4　Learner Comments

At the end of the DDL instruction, reactions about learners' experiences learning grammar with hDDL were solicited through a questionnaire. We grouped learners' comments into the following categories: Effects of DDL, Further Learning, Effects of Group Work, Usefulness and Ease of Use, and Limitations. All questionnaire questions and answers were in Japanese; translations are provided.

Effects of DDL

- ✧ *I have learned from the discoveries of friends. DDL is better than learning from a textbook.*
- ✧ *Since I discovered language rules using hDDL by myself, I felt that I learned more than usual.*
- ✧ *I want to use hDDL to gain a deeper understanding of what I lack in understanding English.*
- ✧ *It was good that I thought for myself instead of being taught by a teacher.*
- ✧ *It was good to find errors in English sentences and look for grammar rules.*

Further Learning

- ✧ *I want to use hDDL and review what I don't understand well.*
- ✧ *I want to look up other grammar items with hDDL.*
- ✧ *I want to use and speak what I learned today.*

Effects of Group Work

- ✧ *I have learned from the discoveries of friends. DDL is better than learning from a textbook.*
- ✧ *It was difficult to find the English rules myself. But after listening to my friends' opinions I could understand.*

Usefulness and Ease of Use

- ✧ *The hDDL is easy to read, so it was easy to understand sentences.*
- ✧ *It was good to look at a lot of example sentences.*
- ✧ *It was good that there were many examples sentences for one grammar point.*

Limitations

- ✧ *DDL takes time to learn.*
- ✧ *It would be better if grammar rules are shown and explained on hDDL.*

5．Conclusion

hDDL was developed to address a lack of level-appropriate corpora and lack of user-friendly tools for learners and teachers. In this study, the procedure for developing such a tool was described, and data from participant questionnaires suggest that it is easy to use. This study, combined with the findings from other studies (Nishigaki, Akasegawa, Kawana et al., 2021; Nishigaki Kawana, Yamaguchi et al., 2021), suggests that hDDL has potential for being an effective corpus tool for lower proficiency level learners, which opens the benefits of DDL to a wider range of learners. It should be noted that these studies have been conducted with small populations, and further studies using non-DDL control groups will provide more insight into the effectiveness and various uses of this tool.

Note

1) The original database and search tool were modified to clear copyright issues, named BES Search and is now available online (https://bessearch.ddl-study.org/). BES Search allows users to search for copyright-free, easy English samples. BES Search includes some unique features. First, a user can customize the search for desired sentences by inputting the length of a sentence, for example, as a "sentence with three to six words," then the search results appear in the order of the sentence length from short to long. Second, BES Search can use Corpus Query Language (CQL) and search for sentences by grammar patterns. In addition, the user can define a search condition using a lemma and/or part of speech.

Acknowledgements

This research was supported by JSPS KAKENHI Grant (B) Number JP20H01277.

References

Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning, 67*(2), 348-393.

Chujo, K., Oghigian, K., & Akasegawa, S. (2015). A corpus and grammatical browsing system for remedial EFL learners. In A. Leńko-Szymańska & A. Boulton, (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 109-128). John Benjamins.

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

Johns, T. (1991). Should you be persuaded: Two examples of data-driven learning. In T. Johns & P. King (Eds.), *Classroom Concordancing. English Language Research Journal, 4*, 1-16.

Leech, J. (1997). Teaching and language corpora: A convergence. In A. Wichmann, S. Fligelstone, A. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 1-23). Longman.

Nishigaki, C., Akasegawa, S., Kawana, T, Nakai, K. Kenmoku, S., & Yamazaki, T.(2021), Classroom application of a web-based DDL support tool in a secondary school, *Proceedings of the JAECS Conference 2021,* 97-102.

Nishigaki, C., Kawana, T., Yamaguchi, A., Orihara, S., Kondo, M., Horne, B. Monoi, N., Hoshino, Y., & Ishii, Y. (2021). Development of DDL support tools for bridging primary, secondary and tertiary vocabulary and grammar learning (handout), Symposium at LET 60th Annual Conference, The Japan Association for Language Education & Technology (LET), (2021, August 21).

Pérez-Paredes, P. (2019). The pedagogic advantage of teenage corpora for secondary school learners. In P. Crosthwaite (Ed.), *Data-Driven learning for the next generation* (pp. 67-87). Routledge.

Römer, U. (2006). Pedagogical applications of corpora: Some reflections on the current scope and a wish list for future developments. Zeitschrift für Anglistik und Amerikanistik [Special issue]. *The Scope and Limits of Corpus Linguistics-Empiricism in the Description and Analysis of English, 54,* 121-134.

Timmis, I. (2015). *Corpus linguistics for ELT: Research and practice*. Routledge.

Wicher, O. (2020). Data-driven learning in the secondary classroom: A critical evaluation from the perspective of foreign language didactics. In P. Crosthwaite (Ed.), *Data-Driven learning for the next generation* (pp. 31-46). Routledge.

Yano, Y., Long, M. H., & Ross, S. (1994). The effects of simplified and elaborated texts on foreign language reading comprehension, *Language Learning, 44*(2), 189-219.