Teaching Remedial Grammar through Data-Driven Learning Using *AntPConc*

* Kiyomi CHUJO
Nihon University
Japan

Kathryn OGHIGIAN
Waseda University
Japan

Laurence ANTHONY
Waseda University
Japan

Kenji YOKOTA
Nihon University
Japan

Abstract

In most Asian countries, students receive between six and eight years of compulsory English education before they enter university. Despite this massive investment in English education, many students, especially in Japan, continue to show a poor understanding of rudimentary grammar rules. In this paper we report on a unique English course designed specifically to address grammar issues at low (remedial) levels using a Data-Driven Learning (DDL) approach. Applications of DDL are becoming more widely reported, but they are generally at the intermediate or advanced level. One of the challenges of using DDL at the remedial level is the lack of suitably leveled corpora. Another challenge is that most corpus tools used in DDL are designed for researchers or advanced learners and thus can appear overly complex. To address these issues, we have developed a simple English corpus built from standard school texts. We have also created a freeware, parallel corpus tool, AntPConc, that is specially designed to be simple, easy, and intuitive to use by beginner learners. Results from the course show significant gains between pre- and post-tests of grammar understanding for beginner-level EFL university students. We also obtained positive student feedback on the *AntPConc* software.

Keywords: Data-Driven Learning; DDL; AntPConc; Concordancer; Remedial Level; Grammar Instruction

Introduction

English education in Japan has been traditionally comprised of six years of study from grade seven when students enter junior high school until they graduate from senior high school at grade twelve. Most university students, however, will have also studied English at private English conversation schools and/or cram schools. As part of major education reform in 2011, the Japanese Ministry of Education adjusted this system and introduced a further two years of compulsory education at grades five and six of elementary school. This has brought English schooling in Japan more in line with that of other Asian countries, including Taiwan, China and Korea.

Despite a huge investment in English schooling in Asian countries, both scholars and major media outlets have continued to report on the poor English level of students, especially in Japan. Ono et al. (2005), for example, reported that many Japanese university students have a poor understanding of even basic grammar rules. Chujo et al. (2012) administered a basic grammar test to 160 university Japanese freshmen and found that they failed to answer correctly almost 30% of junior high school grammar items and 55% of senior high school grammar items. These results are consistent with those of Manaka (2010) and Zhang and Hirasawa (2010) who describe the English at many Japanese universities as being at the 'remedial' level. It is perhaps because of this poor English level that Japanese student scores on international tests of English, such as TOEIC and TOEFL, are among the lowest in Asia.

In recent years, a growing number of studies have reported on the effectiveness of Data-Driven Learning (DDL) in both general English and ESP classroom settings (Lee & Swales, 2006; Anthony, 2011; Boulton, 2012; Geist & Hahn, 2012; Osolsobě & Vališová, 2012). DDL can be defined as an approach to language learning that encourages the use of authentic materials, promotes learner-centeredness through real exploratory tasks and activities, and makes effective use of corpus tools, such as concordancers. To date, many implementations of the DDL approach have been at the intermediate or advanced level (Boulton, 2008). Weber (2001), on the other hand, has suggested that the approach might also be effective with remedial students, and Takanashi (2009) has suggested that DDL might even be useful for introducing beginner level grammar.

In order to implement DDL in a beginner-level classroom, two challenges need to be overcome. First, there is a need for a corpus of target language at an appropriate level for the students. Traditional corpora of authentic native speaker language are simply far too difficult. Second, there is a need for a simple and easy-to-use corpus tool that will not overwhelm the students and prevent them from focusing on the task at hand. Unfortunately, the majority of currently available corpus tools are designed for researchers or advanced learners and are thus overly complex for the beginner-level classroom.

In this paper, we describe a novel ESP course for university students designed specifically to address grammar issues at low (remedial) levels using a DDL approach. To meet the challenges described above, we have developed an English corpus built from traditional school texts that highlights target grammar items. We have also created a freeware, parallel corpus tool, *AntPConc*, that is specially designed to be simple, easy, and intuitive to use by beginner learners. Results of pre- and post-tests of grammar understanding show that students are able to make significant gains over the duration of the course. We also obtained positive student feedback on the AntPConc software, suggesting that it can be applied in a wide-range of learning environments. In the following sections, we will first describe the specialized grammar corpus and the new *AntPConc* tool that we developed. Next, we will explain the grammar-based syllabus of the course and the four-step methodology that we employed. Finally, we will make our conclusion and suggest areas for further study.

Development of a Specialized Grammar Corpus and the AntPConc Corpus Tool A parallel corpus of beginner level English (CoBLE)

A central component of any DDL course is a corpus of target language. For our goal, we required a corpus that would allow students to investigate the usage of beginner level grammar items but at the same time not overwhelm them with colloquial or low-frequency vocabulary. We also required a full-text corpus so that we could present students with target items in complete-sentence contexts. Further, we required that the sentences in the corpus would be short and simple, ideally fitting on a single line in a concordancing tool.

Two of the most popular corpora used in corpus studies are the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA). However, we were not able to capitalize on these resources because both would expose students to completely authentic, native speaker language that included vocabulary that would be far beyond the learners' levels. Also, in the case of COCA, the interface to the corpus provides only partial sentences, rendering it even less useful.

Due to the lack of a suitable publicly available corpus, we compiled an original corpus specifically for our target students. In an earlier study, we identified several resources that were appropriate for this target population including English school textbooks used in Asian countries, American reading textbooks, and English graded readers (Chujo, Nishigaki, Yamaho & Ochiai, 2012). Based on this information, we first collected 10,352 English sentences (66,104 English word tokens) from American language textbooks for grades one to six. Then, to provide further support for students in the DDL classroom, we created a Japanese translation for each sentence, using machine translation followed by manual correction. An investigation of the final Corpus of Beginner Level English (CoBLE) revealed that it covers almost all of the targeted grammar items at the junior high school level.

AntPConc: a freeware parallel concordancer

For the successful implementation of a DDL course, students need to be able to easily and intuitively operate a concordancer, i.e., a corpus analysis software tool. For our purpose, we also required a concordancer that could process our parallel corpus of English target texts and their Japanese translations. In addition, we required a tool that students could use not only in the classroom as part of regular exercises but also outside of the classroom as part of homework.

Previously, the only available parallel corpus tool was *ParaConc* (Barlow, 2004), which is a CD-ROM-based commercial software tool designed for use by translators and corpus researchers. Although *ParaConc* was used in earlier studies (see Chujo & Oghigian, 2008, 2012), it was found that this tool required training before use, needed default settings changed at each launch, and limited access to only

classroom use due to copyright reasons. Therefore, we decided to build a freeware, standalone, parallel concordancer, *AntPConc*, and released a prototype version to the public in 2012 (Anthony, 2012). Figure 1 shows a screenshot of *AntPConc* after our beginner-level parallel corpus has been loaded and a search has been made. As can be seen in this screenshot, *AntPConc* has a clear and very minimal interface, so that students are able to use it independently to explore grammar patterns in the concordance lines. All aspects of the design were made with multimedia principles in mind, that is, to allow the user to easily search and sort grammatical features in a language and to compare them with those in another language. Users simply type the search term into the box (indicated with a red circle) and press *start*, and the results are shown in full sentences using color-coding to highlight the pattern. Additionally, it can be used not only with English and Japanese parallel corpora, but also with corpora of any other language, including Chinese and Korean. Furthermore, it can display three or more corpora of different languages in the same interface provided they are encoded in the international UTF-8 character encoding standard and are aligned minimally at the file-line level.

Figure 1. *Screenshot of AntPConc*
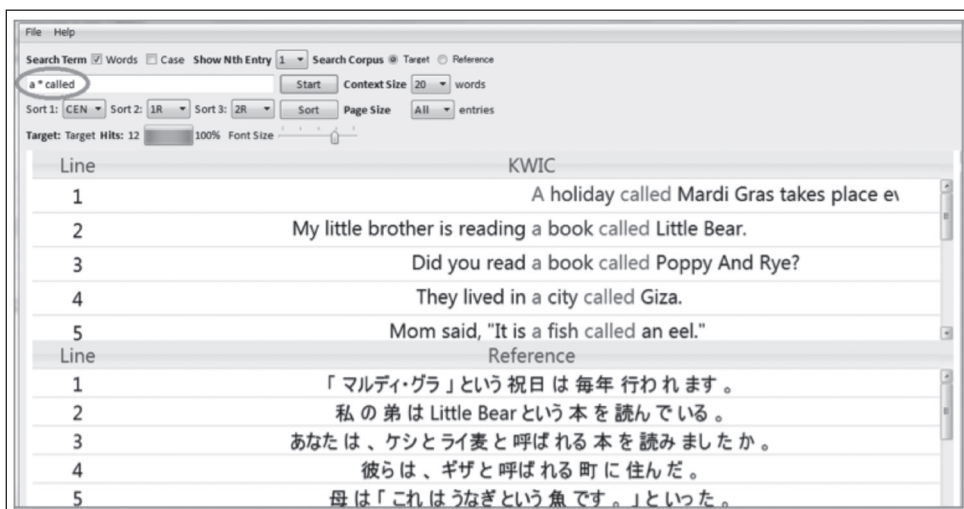
Although it is always difficult to know at the design stage if a software tool will be easy and intuitive to use by the end user, a number of standard design principles were followed in the development of *AntPConc*. First, the software was designed according to the Microsoft Experience Interaction Guidelines (Microsoft, 2013), an 800 page document describing all aspects of interface design from basic design rules to the positioning of interface widgets (e.g., menus, search boxes, check buttons, scrollbars, etc.). In addition, the basic common operations that we expected users to carry out frequently, such as switching languages, changing the sort order, and activating case options, were made accessible directly on the main screen. Lonfils and VanParys (2001) explain that this is an essential feature of good software design as it avoids the need for additional pop-up windows or confusing pull-down menus. We also gave the tool's interface widgets the native look and feel of the operating system, which Lonfils and VanParys explain will improve the ease-of-use and intuitive operation of the program as it will match with the learners' habits and expectations. We also engaged in an extensive testing program, whereby beta versions of the tool were trialed by teachers of the course with feedback from them leading directly to changes to the tool itself.

Teaching Grammar to Remedial Students Using DDL

In our ESP course for remedial English grammar, the participants were twenty-two first-year university engineering students (19 males and 3 females) with an average TOEIC Bridge score of 118. Our goal in this study was to assess if students could improve their understanding of remedial grammar through the DDL approach using our CoBLE parallel corpus of beginner level English together with *AntPConc*. The study was carried out in 2012 and ran over two semesters for a total of 15 weeks (five weeks in the spring semester and ten weeks in the fall semester). Each week, students participated in a 45-minute DDL session. The learning effect was measured by the gain between a pre-test and post-test, and we collected student feedback to get a sense of how students felt about the approach.

Grammar-based syllabus

In order to develop an appropriate grammar-based syllabus for our students, it was first necessary to identify their specific grammar needs. To do this, we employed the basic grammar proficiency test developed by Chujo et al. (2012). This test resulted from an investigation of the English proficiency levels of elementary and junior high school students carried out by Shirahata (2008) and an investigation of specific grammar weaknesses of high school students taking the TOEIC test carried out by Uchibori, Chujo and Hasegawa (2006). For our study, test items that were incorrectly answered by almost 30% of the university students were selected for inclusion in the syllabus and later targeted in language exercises.

Table 1 shows the resulting grammar syllabus that we employed. The DDL lessons began from the sixth week of the spring semester when the CoBLE parallel corpus was ready for use. The first items covered were high-error items that were easy for us to work into DDL tasks. Following these, items with lower error rates on the proficiency test were introduced. However, the order of items is likely to be revised in future years depending on the analysis of the results and student feedback from this current study.

Table 1. *Remedial Syllabus Based on Junior High School Grammar Items*

Week	Spring Semester	Fall Semester
1		Comparison
2		Modal auxiliaries (may, shall, must)
3		Present perfect
4.		Yes-no questions (do, does, did, was, were, …)
5	(Pre-test)	Tense (present, present progressive, past, …)
6	Nouns (countable & uncountable nouns)	Pronouns (possessive pronouns)
7	Adjectives (participial adjectives)	Conjunctions (if and when)
8	Nouns (possessive nouns)	Relative pronouns (who, which, that)
9	Wh-questions (what, what time…)	Indirect questions (Do you know where…)
10	Existential there (there is, there are)	Special uses of it (time & weather)
11	(Term-end test)	(Post-test)

Four-step grammar teaching procedure

Table 2 shows the four-step procedure developed by Chujo and Oghigian (2008) that the students followed. In Step 1, students explored six tasks in a specific grammar context using the DDL approach and employing *AntPConc*. For these tasks, students generally worked in pairs sharing their discoveries and offering each other support, eventually arriving at hypotheses about the form and usage of a particular grammar pattern. In Step 2, the teacher explained the grammar items so that students could confirm, clarify or correct the hypotheses they made through the inductive DDL tasks. In Step 3, students completed additional practice and consolidation exercises individually for homework. Finally, in Step 4, students worked together to complete the production practice exercises in class.

As can be seen, the procedure is similar to a traditional (non-DDL) class procedure in which the teacher first presents a grammar point or language feature, then explains it explicitly, and finally asks students to practice and produce it. However, whereas students are often passive in the first step of the traditional approach, in the first step of our approach, it is the teacher that takes a more passive role. In this step, the teacher allows students to freely observe and notice grammatical patterns while they form hypotheses about what they are seeing. The teacher's role, therefore, is to encourage implicit discovery learning, i.e., to make sure students actively think about what they are seeing and why.

In the second step of our approach, the teacher presents both a written and verbal explanation of the grammar pattern in Japanese. This allows students to confirm, clarify or correct their own understanding. It also ensures that all students are looking at and understanding the same thing, which is especially important as they may have observed different phenomena and created divergent hypotheses at stage one. In other words, here, the teacher's role is to ensure that all students are "on the same page." Each year, students participating in this type of DDL course have reported in evaluation surveys that they particularly appreciated the explicit explanation that was given after they had an opportunity to explore and discover the targeted grammar. They explained that this helped them to clarify their understanding in a way that was

familiar to them, and providing it after rather than before the discovery learning step allowed them to reach a deeper understanding of the language.

Another important difference compared to a traditional class procedure is the role of peer support. In stage one of our approach, students sit in pairs at computers following a worksheet and examining the *AntPConc* concordance display for particular information or patterns. Here, they are encouraged to work with each other discussing how to perform the concordance search and how to interpret the patterns and tendencies in the lines they found. This pair and group discovery learning has a number of advantages over individual work. The first is that it is possible to avoid having a lone student who is lost and silent. Secondly, students can discuss and debate their hypotheses with others creating more refined hypotheses. Thirdly, if further clarification is needed, the students are more likely to ask each other rather than interrupt the teacher. Finally, since the students have already worked together on the exploration of the corpora, they will be at a better comfort level when they need to work in pairs and groups to practice and produce language patterns at stage four.

Table 2. *Four-Step Grammar Teaching Procedure*

Step 1	Hypothesis formation through inductive DDL tasks with hands-on AntPConc
Step 2	Explicit explanations from the teacher to confirm or correct these hypotheses
Step 3	Hypothesis testing through follow-up exercises (homework) and teacher feedback on homework
Step 4	Production through follow-up exercises (in class) and teacher feedback on homework

Figure 2 shows an example guided exercise from Step 1 of the grammar teaching procedure. The aim of the exercise is to develop an understanding of countable and uncountable nouns. In the figure, only the English lines, without the Japanese translations, are shown due to the space limitations of the present article. Also, the answers below the display are provided for the reader and are not included in the original student version.

Figure 2. *Example AntPConc DDL Exercise 1*

Search some. Write down "some + noun" and underline it if there is no "s" at the end.

```
          We need some help.
      Dana eats some food.
       Lee ate some popcorn.
          Draw some squares.
      Here are some blue dishes.
      We read some books.
      We find some wood.
     Jen saw some frogs on the rocks.
       I want some milk.
   Annie added some soap.
          Pack some food.
 Anna asked, "May I have some peanuts, Miss Rios?"
   Take along some water.
            Do some groups of stars look like animals?
 She played on some swings.
```

Answers:

   some help, some food, some popcorn, some squares, some blue dishes, some books,

   some wood, some frogs on the rocks, some milk, some soap, some food, some

   peanuts, some water, some groups of stars, some swings

In this exercise, students, working in pairs, type in the word "*some*" and look to see both "*some* + countable nouns" and "*some* + uncountable nouns" patterns. After writing these down, they observe what comes after "*some*." When they find "*some* + uncountable nouns" they underline them. By following this procedure, students learn to identify the head of a noun phrase (here, *some* + noun) by concentrating on the quantifier *some*, which is responsible for determining the lexical properties designated by the noun phrase (e.g., *some books, some food*). In other words, the quantifier *some* heads the noun phrase, wherein *some* has the function of "instantiating" a "type" of thing (i.e., an individual object or a substance). Usually, a total of six exercises related to the unit goal are given.

After completing the stage one exercises, students summarize what they have learned through the DDL activities at the bottom of their worksheets. These worksheets are collected by the teacher and quickly checked to see if the students

have attained the goal of the lesson. If necessary, the teacher answers the students' questions and adds corrections. The students' reflective summaries of the grammar points provide students with a way to verify their hypotheses and provide the teacher with a way to confirm if the main patterns focused on in the class have been noticed.

Figure 3 shows another guided exercise from Step 1 of the grammar teaching procedure. Here, the aim of the exercise is to develop an understanding of the role of possessive apostrophes. First, students type in apostrophe -'s and then sort the output. They are then asked to guess what type of word correctly completes the blank in each sentence (the answers are shown in parentheses), and they find the exact answers from the *AntPConc* concordance line display. Next, they write down a sentence including the answer. Students are also asked to write down each translation and then look for the corresponding Japanese pattern. By following this procedure, students learn to identify apostrophe -'s as the head that selects for the complement of a complex noun (e.g., *boy's*), and to produce a larger phrase (e.g., a *boy's bedroom*).

Figure 4 shows a third example of a guided exercise. Here, the aim of the exercise is to develop an understanding of regular and irregular past verb forms. First, students type in the target adverbs of time (e.g., yesterday). Next, they are asked to write down both the English and Japanese sentences that are displayed and underline the adverbs and the corresponding verbs. In Figure 4, students can find that *yesterday* is used with past verbs, that simple past verbs such as *arrived, finished*, and *visited* are formed by adding *-ed* or *-d*, and that some verbs like *got* and *fell* have irregular past forms. From the translation they can also find the corresponding Japanese pattern Verb-(*shi*)*ta* 'Verb Past'. By following this procedure, students can be encouraged to make a conscious effort to use regular and irregular past verbs forms correctly.

Figure 3. *Example AntPConc DDL Exercise 2*

Search 's. Fill in the blanks in the table below and write a sentence which includes 's from sentences you found.

```
            We need some help.
         Dana eats some food.
         Lee ate some popcorn.
            Draw some squares.
       Here are some blue dishes.
         We read some books.
         We find some wood.
        Jen saw some frogs on the rocks.
          I want some milk.
      Annie added some soap.
            Pack some food.
  Anna asked, "May I have some peanuts, Miss Rios?"
      Take along some water.
              Do some groups of stars look like animals?
     She played on some swings.
```

Answers:

boy	(boy's)	boys	boys'	It was a boy's bedroom. （少年の寝室）
brother	(brother's)	brothers	brothers'	My brother's friend asks about my brother. （兄の友人）
friend	(friend's)	friends	friends'	My friend's father said, "Young man, go home."（友人の父）
man	man's	men	(men's)	The men's faces were very white.（男の顔）
child	child's	children	(children's)	Children's shoes have to be tough. （子供達の靴）

Figure 4. *Example AntPConc DDL Exercise 3*

Search yesterday. Write down 5 sentences with the translation. Then underline verbs and adverbs of time.

```
            We arrived in Brazil yesterday.
            Josh got a new bike yesterday.
       I finished a book about birds yesterday.
   Today we practiced longer than we did yesterday.
   I painted more badly today than I did yesterday.
            The first snow fell yesterday.
```

```
         私 たち は 昨日 Brazil に 到着 した 。
         Josh は 昨日 新しい 自転車 を 買った 。
       私 は 昨日 鳥類 について の 本 を 読み 終えた 。
         今日 、 我々 は 昨日 より も 長く 練習 した 。
   私 は 昨日 より 今日 の 方 が もっと ひどい 絵 を 描いた 。
            昨日 、 初雪 だった 。
```

Answers:

We **arrived** in Brazil **yesterday**.	私たちは**昨日**ブラジルに**到着した**。
Josh **got** a new bike **yesterday**.	ジョシュは**昨日**新しい自　車を**買った**。
I **finished** a book about birds **yesterday**.	私は**昨日**鳥類についての本を＿**み終えた**。
Sally and Chip **visited** us **yesterday**.	サリーとチップは**昨日**私たちを**訪問した**。
The first snow **fell yesterday**.	**昨日**，初雪**だった**。

Figure 5 shows an example practice and consolidation exercise from Step 3 and Step 4 of the grammar teaching procedure. In Step 3, students practice and consolidate their hypothesis testing skills by completing tasks for homework, such as Tasks 1 and 2 in the figure. In Step 4, students complete production tasks, such as those in Tasks 3 and 4. The answers here are provided for the reader and are not included in the student version. Students are given approximately twenty-five questions as homework. Since class time is limited and there is only so much homework that a teacher can assign, we are now investigating ways to increase production during class time.

Students Gains

Pre- and post-tests

Table 3 shows sample questions and answers we used as part of the pre- and post-testing. Students filled in the blanks by looking at each Japanese translation. The test had 34 questions which were based on the same targeted grammar items shown in the syllabus in Table 1. These test items are based on two foundational studies on grammar items taught in Japanese secondary schools (Shirahata, 2008; Uchibori et al, 2006) and have been used in several case studies. Cronbach's alpha for the 34 items is 0.8, suggesting that the items have relatively high internal consistency. We used the same test items for the pre- and post-test, but the students were not told they would be given tests, and they were not given the answers at any time.

Figure 5. *Sample Follow-up Tasks*

Homework

Practice & Consolidation

Task 1 Choose the best answer.

* Mark (lives, <u>has lived</u>) in Canada since April.

* Jane and I are friends. (I know,<u> I've known</u>) her for a long time.

Task 2 Write the simple past and the past participles.

* buy   (bought) (bought)

* write   (wrote)   (written)

Production

Task 3 Write sentences about the past.

* Joey often loses his keys. (He lost his keys) last week.

* I eat an apple every day. Yesterday (I ate an apple).

Task 4 Complete each English sentence.

* Do you know (where the rest room is)? (トイレがどこか知っていますか ?)

* Do you know (who wrote Snow Country)? (だれが雪国を書いたか知って いますか ?)

Table 3. *Pre- & Post-Test Sample Questions*

1. 何人かの　生は電車で　校に　る。   Some (students) come to school by train.
2. これは私の先生の家です。   This is my (teacher's) house.
3. ドアをあけましょうか？   (Shall) I open the door?
4. 明日何をするつもりですか？。   What (are) you (going) (to) do tomorrow?
5. 富士山は日本で最も高い山です。   Mt. Fuji is (the) (highest) mountain in Japan.
6. 私は宿題を終えたところです。   I (have) just (finished/done) my homework.

Results from the re- and post-tests

    Table 4 shows the descriptive statistics such as average percentage scores of correct answers and the standard deviation (SD) of the average percentage scores of correct answers. The class gained an average of 8.29 points, from 76.47% to 84.76%. Since the skewness and kurtosis were not within ±2.0, showing the lack of symmetry in the distribution of the data value, a non-parametric test, the Wilcoxon signed-rank test, was applied to the pre- and post-test scores. The results are shown in Table 5 and Table 6, where it can be seen that there was a significant increase with a difference significant at the 1% level, with $z = 3.70$, and $p < .001$. The effect size was shown to be large (Pearson's correlation coefficient $r = .79$) and thus we can say that students improved their remedial grammar skills as measured by such a test. In 2012, we taught only junior high school remedial grammar items, which were rather easy for some students. Therefore, in a future study, we hope to focus on grammar items identified as difficult for senior high school students. It should be noted that although a pre-test score of 76.47% may seem to be a reasonable reflection of learning, the targeted items are very simple grammatical features that students theoretically would have understood at a 100% level while in junior high school. In addition, the average TOEIC Bridge score for these students was 118, which is equivalent to the Eiken Grade 4 level (used often in Japan) and as such, has been identified by the Japanese Ministry of Education, Culture, Sports, Science and Technology as below the stated goal for the English proficiency level for junior high school students (MEXT, 2002).

Table 4. *Descriptive Statistics*

	N	Mean	SD	Minimum	Maximum	Skewness	Kurtosis
Pre-test	22	76.47	14.12	32.35	91.18	3.25	2.56
Post-test	22	84.76	10.05	55.88	94.12	-1.60	-1.74

Table 5. *Ranks*

		N	Mean Rank	Sum of Ranks
Post - Pre	Negative Ranks	1a	3.00	3
	Positive Ranks	18b	10.39	187
	Ties	3c		
	Total	22		

a.   Post < Pre b. Post > Pre c. Post = Pre

Table 6. *Test Statistics*

	Post - Pre
Z	3.7023
Asymp. Sig. (2-tailed)	.0002

Although control groups were not used (for ethical reasons, we wanted all students to have access to this new technology), Chujo et al. (2013) used a meta-analysis on the pre- and post-test design for nine similar DDL studies preceding this study and obtained the combined effect size, as measured by Cohen's d, of 1.26 and the 95% confidence intervals of 1.05 to 1.47. (In this current study, Pearson's correlation coefficient r was calculated.) An effect size of 1.26 is very high, indicating that DDL can be effective by conventional estimates.

Table 7 shows some of the grammar items and example questions on which students made excellent gains in the December post-test. The number shows the percentage of students who gave correct answers. These results seem to indicate that providing multiple examples in the sentence-based concordance lines together with their translation was effective in helping students to learn these items.

Table 7. *Items Showing Highest Gains*

Grammar items and questions	Pre-test (April)	Post-test (December)
Past tense (irregular verbs) Mrs. Suzuki (gave) this cap to me.	55%	86%
Possessive pronouns This blue book is Yuki's, but that red one isn't (hers).	68%	91%
Possessive nouns This is my (teacher's) house.	50%	86%
Was …? (Was) Taro an elementary school boy five years ago?	68%	91%
Be going to …? What (are) you (going) (to) do tomorrow?	55%	86%

Table 8 shows two grammar items and example questions for which students initially showed gains (after five weeks), but later showed no improvement (after 15 weeks). These items are *there is/are* and the plural forms of regular nouns (*-s*). When learning another language, comparisons are often made to the student's L1 (Japanese). In many cases, a new L2 (English) feature generally can be identified in one form or another in the L1. In the Japanese language, something either exists or does not exist and a pronoun (*there*) is not used. Students therefore need to make a conscious effort to use the existential construction, which contains the dummy pronoun *there*, the verb be, and the displaced subject, alone or accompanied by optional phrases like *in my room*. The lack of this syntactic means of expressing existential propositions in their language may not draw the learner's attention readily to the combination of *there* and *be*.

The analysis based on the head/non-head identification in *Cognitive Grammar* (Taylor, 2002) applies quite generally to the *there* construction. The fragment, *there be*, should be identified as a non-head (modifier), since the immediately following string, *a desk in my room*, basically represents a proposition. Although this fragment lacks an overt predicate, students may have successfully arrived at the interpretation that *a desk is in my room*. This would allow students to obtain a reasonable interpretation of this grammatically incomplete fragment without the need for the existential phrase there is. From this interpretation, students may not have been able to retain the existential expression easily.

Regarding the second example sentence in Table 8, the plural -*s*, attached to the base noun *student* can be conceived of as a head that designates a plural number of entities (here, *students*). In *Cognitive Grammar*, an affix functions as a head taking a noun as its argument. However, for some learners, the plural -*s* arguably did not function as a head, but as a non-head. This incorrect identification may have, again, prevented students from retaining the regular plural nouns easily.

Table 8. *Items Showing Initial Gains That Are Later Lost*

Grammar items and questions	April	July	December
There is/are (There) (is) a desk in my room.	68%	91%	68%
Plural forms of nouns (regular) Some (students) come to school by train.	77%	91%	77%

AntPConc: Student Feedback

Student feedback on AntPConc

All 22 students in the study had no prior experience using corpora. At the end of the fall semester, we asked them to rate the AntPConc corpus tool according to five items (shown in Table 9) using a five-point Likert rating scale from "strongly agree" (rating 5) to "strongly disagree" (rating 1). We also asked them to make suggestions for further improvements and offer general comments on their experience using the software. Selected responses translated from Japanese are shown in Table 10. Results showed that students rated the tool quite highly for accessibility, intuitive user interface, and screen design. Speed received a reasonable mean and very few students experienced problems. Only three students had slight problems such as adding an unnecessary space before a wild card, or being unable to load both English and Japanese corpus files easily. Also, one student forgot to click the SORT button after generating the initial results, leading to discrepancies between the printed exercise instructions and the actual display.

These results suggest that students were generally satisfied with AntPConc. Also, from the teacher's view point, students quickly learned how to use the software in just

the first lesson and were able to then concentrate on the DDL tasks in the remaining classes. The teacher needed only to direct them on the worksheet with instructions such as "click the File button, then click the Load Corpus Files button, then click the Choose Files button, and then open the English text and Japanese text." Using AntPConc in the classroom was found to be very easy, and no teaching assistants or introductory sessions were necessary. The teacher also reported being much more relaxed than in previous computer-based DDL classes. Of course, software design should always be considered to be an evolutionary process, and so the students' suggestions for improvements are now being considered as the next version of the software is being developed.

Table 9. *Student Feedback on AntPConc*

1 (strongly disagree) --- 5 (strongly agree)	Mean	SD
Easy accessibility	4.6	0.7
Intuitive user interface	4.3	1.0
Appealing screen design	4.3	0.9
Fast speed	4.1	1.1
I experienced problems	2.0	1.4

Table 10. *Student Comments on AntPConc*

Selected positive comments
I was able to use it intuitively.
Instantly I was able to search many words.
The software was simple and user-friendly.
I could easily look at ample examples of various usages.
When I clicked a sentence, the blue highlighting was useful.
Selected negative comments
The letters were too small to read.
Please leave a little bit more space between lines.
Sometimes the speed was not fast enough.
I didn't understand the SORT well.
The translation was sometimes too direct.

The role of L1 translation

We also asked the students if they thought the L1 (Japanese) translation in the parallel corpus was necessary using a five-point Likert rating scale from "strongly agree" (rating 5) to "strongly disagree" (rating 1). The results are shown in Table 11. If we combine the agree and strongly agree responses, we see that 20 out of 22 students (90%) found the translations to be necessary. Interestingly, this rate is much higher than the 75% rate of beginner level students from previous studies who used DDL for noun and verb phrases (Chujo & Oghigian, 2012: 179) and who had slightly higher English proficiency levels. Thus, these results suggest that the translation may allow learners to focus on the grammatical structure without struggling with the meaning. Furthermore, as the results in Table 7 show, some grammatical items, such as possessive pronouns and nouns, have corresponding Japanese expressions and so the translation can draw the learner's attention readily to the item. However, we should note that we see the Japanese translation as a supportive measure that students will use less and less over the course of the semester, but are nevertheless glad to have.

Table 11. *The Role of L1 Translation*

It is necessary for me.			n = 22	
1 *(strongly disagree)*	2 (disagree)	3 (neutral)	4 (agree)	5 (strongly agree)
0 (0%)	0 (0%)	(9%)	(18%)	(72%)

Conclusions

Previous studies have shown that a dismaying number of incoming university students lack a basic understanding of grammar, especially in Japan. To address this problem, we have proposed a new and effective way to engage remedial level students in understanding language, with a very practical focus on the grammatical items they were taught in secondary school, but failed to acquire. Our approach is based on the concept of Data-Driven Learning (DDL) and utilizes the CoBLE beginner-level

English corpus that is specifically designed to target remedial-level grammar items. Our approach also capitalizes on an easy-to-use, freeware parallel corpus analysis tool, AntPConc, designed specifically for use by novice users of computers.

Results from our study show that (1) DDL can be effective for remedial level university students in teaching basic grammar; (2) students can make significant gains using the CoBLE corpus together with *AntPConc*; and (3) both students and the teacher respond well to this new standalone parallel concordancer. Future studies will focus on refining the DDL tasks and expanding the syllabus to include targeted senior high school grammar items to improve remedial students' basic grammar skills. We are also planning to improve the usability of AntPConc further by allowing multiple corpora to be loaded into the system with a single click, introducing a combined search and sort function, and offering further display options, such as user-defined font sizes and color preferences.

This study is a follow-up to a series of studies conducted annually since 2004 to investigate the usefulness of DDL in an L2 classroom (see Chujo & Oghigian, 2008; Chujo & Oghigian, 2012). Each year, participants have provided positive feedback on DDL in terms of its novelty, its ability to help students to concentrate on their work, and its usefulness for learning grammar and vocabulary in an easy to understand way. Although modifications and improvements have been made with each subsequent study, for example in 2013 we introduced *AntPConc*, we have not been able to follow participants longitudinally to track their continued use of corpora or their long-term retention. As a result, we cannot say that students continue to use DDL after the course has finished, nor can we confirm that their implicit learning stays with them. Similarly, not all studies have been conducted with control groups, although we are now collecting data using a DDL class and a control group (a separate but similar class) as an extension of this study using *AntPConc*. We can, however, confirm that students have consistently made gains during the year they participated in these studies. In addition, research using meta-analysis of numerous DDL studies (Cobb & Boulton, in press and Chujo et al., 2013) has provided clear evidence on the effectiveness of DDL in the L2 classroom.

Acknowledgments:

References

Anthony, L. (2011). Introducing corpus-based methods into a large-scale technical writing program for scientists and engineers. *Proceedings of the Corpus Linguistics Conference (CL2011)*. Retrieved November 4, 2013 from http://www.birmingham.ac.uk/documents/college-artslaw/corpus/conference-archives/2011/abs-263.pdf.

Anthony, L. (2012). AntPConc (Version 0.05) [Computer Software]. Tokyo, Japan: Waseda University. Available from http://www.antlab.sci.waseda.ac.jp/

Anthony, L., Chujo K. & Oghigian, K. (2011). A novel, web-based, parallel concordancer for use in the ESL/EFL classroom. In J. Newman, H. Baayen & S. Rice, (Eds.), *Corpus-based Studies in Language Use, Language Learning, and Language Documentation* (pp. 123-138). Amsterdam/New York: Rodopi Press.

Barlow, M. (2004). Paraconc [Computer software]. Houston: Athelstan. Available at http://www.athel.com/mono.html

Boulton, A. (2008, March). Evaluating corpus use in language learning: State of play and future directions. Paper presented at the American Association of Corpus Linguistics, Provo, U.S.

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

Chujo, K. & Oghigian, K. (2008). A DDL approach to learning noun and verb phrases in the beginner level EFL classroom. In A. Frankenberg-Garcia (Ed.) *Proceedings of the 8th Teaching and Language Corpora Conference*, (pp. 65-71). Lisbon, Portugal: Associação de Estudos e de Investigação Científica do ISLA-Lisbo Fichier.

Chujo, K., Nishigaki, C., Yamaho, M. & Ochiai, T. (2012). Beikoku reading kyoukahso to Eigo graded readers no Eigo shokyuusha-muke data toshiteno tekisei ni kansuru kenkyuu [Identifying the suitability of American reading textbooks and English graded readers for beginner-level corpus data]. *Journal of the College of Industrial Technology, Nihon University, 45*, 29-42.

Chujo, K. & Oghigian, K. (2012). DDL for EFL beginners: Recent gains and student views on the role of L1 and paper-based concordancing. In J. Thomas & A. Boulton (Eds.), *Input, Process and Product: Developments in Teaching and Language Corpora* (pp. 130-143). Brno, Czech Republic: Masaryk University Press.

Chujo, K., Yokota, K., Hasegawa, S. & Nishigaki, C. (2012). Remedial gakushuusha no Eigo shuujukudo to Eigo bunpou jyukutatsudo chousa [Identifying the general English proficiency and distinct grammar proficiency of remedial learners]. *Journal of the College of Industrial Technology, Nihon University, 45*, 43-54.

Chujo, K., Anthony, L., Akasegawa, S., Nishigaki, C., Mizumoto, A. & Utiyama, M. (2013, October). Kyouiku riyou kanouna parallel corpus kensaku platform no kouchiku ni mukete [Toward building a parallel corpus concordance platform for English language education]. Paper presented at the 39th Japan Association for English Corpus Studies, Sendai, Japan.

Cobb, T. & Boulton, A. (in press). Data-driven learning: Definition, rationale, and outcomes. In D. Biber & R. Reppen (Eds.), *Classroom Applications of Corpus Analysis Cambridge* Handbook of Corpus Linguistics.

Geist, M. & Hahn, A. (2012). Using a corpus for written production: A classroom study. In J. Thomas & A. Boulton (Eds.), *Input, Process and Product: Developments in Teaching and Language Corpora* (pp. 123-135). Brno, Czech Republic: Masaryk University Press.

Green, B. (2006). A framework for teaching grammar to Japanese learners in an intensive English program. *The Language Teacher, 30*(2), 3-11.

Lee, D. & Swales, J. (2006). A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora. *English for Specific Purposes*, 25, 56-75.

Lonfils, C. & Vanparys, J. (2001). How to design user-friendly CALL interfaces. *Computer Assisted Language Learning, 14*(5), 405-417.

Manaka, W. (2010). Kiso level no daigakusei ni chuugakusei wo shidou saseru kokoromi [A case study of college remedial students teaching to junior high school students]. *Journal of the Japan Association for Developmental Education, 5*(1), 21-27.

Microsoft. (2013). User Experience Interaction Guidelines. Retrieved November 4, 2013 from http://msdn.microsoft.com/en-us/library/aa511258.aspx

Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2002). Eigo-ga tsukaeru Nihonijn no ikusei no tameno senryaku kousou no sakutei nit suite [A strategic plan to cultivate Japanese with English abilities]. Tokyo, Japan. Available at http://www.mext.go.jp/b_menu/shingi/chousa/shotou/020/sesaku/020702.htm.

Ono, H., Muraki, E., Hayashi, N., Sugimori, N., Nozaki, H., Nishimori,…Sakai, S. (2005). Nihon no daigakusei no kiso gakuryo ku kouzou to remedial kyouiku [A development of placement test and e-learning system for Japanese university students: Research on support improving academic ability based on IT], *NIME Research Report 6*, 1-147.

Osolsobě, K. & Vališová, P. (2012). Using data-driven methods in teaching Czech as a foreign language. In J. Thomas & A. Boulton (Eds.), *Input, Process and Product: Developments in Teaching and Language Corpora* (pp. 183-194). Brno, Czech Republic: Masaryk University Press.

Shirahata, T. (2008). Shougakusei to chuugakusei no eigo jyukutatsu do chousa [The i nvestigation into the proficiency levels of English by both elementary school children and junior high school students]. In I. Koike (Ed.), *Daini gengo shuutoku kennkyuu wo kiban to suru shou, chuu, kou,dai no renkei wo hakaru eigo kyouiku no sendou-teki kennkyuu [Leading research to coordinate primary, secondary, and tertiary level English education based on second language research]* (pp. 166-190). *Kaken Kenkyuu Seika Houkokush.*

Takanashi, Y. (2009). *Data de yomu Eigokyouiku no Joushiki [Understanding Common Views about English Education from Data]*. Tokyo: Kenkyuusha.

Taylor, J. R. (2002). *Cognitive Grammar*, Oxford: Oxford University Press.

Uchibori, A., Chujo, K. & Hasegawa, S. (2006). Towards better grammar instruction: Bridging the gap between high school textbooks and TOEIC, *Asian EFL Journal, 8*(2), 228-253.

Weber, J. (2001). A concordance- and genre-informed approach to ESP essay writing. *ELT Journal, 55*(1), 14-20.

Zhang, R. & Hirasawa, K. (2010). Eigo remedial kyouiku no class wake houhou ni tsuite [Student classification for English remedial education], *Journal of the Japan Association for Developmental Education, 5*(1), 60-66.