

## **Research**

# **Paper-Based, Computer-Based, and Combined Data-Driven Learning Using a Web-Based Concordancer**

**Kiyomi Chujo**  
Nihon University, Japan

**Laurence Anthony**  
Waseda University, Japan

**Kathryn Oghigian**  
Waseda University, Japan

**Asako Uchibori**  
Nihon University, Japan

### **Abstract**

Very few studies have investigated the use of data-driven learning (DDL) in the beginner-level EFL classroom, and few or no studies have compared the use of paper-based, computer-based and combined approaches. This paper reports on the results of a three-year comparative case study of computer-based, paper-based, and combined computer- and paper-based DDL using a parallel corpus for beginner-level university students. Students followed guided tasks on a worksheet to inductively understand target grammar patterns, had an explicit confirmation or correction of their hypotheses, and did follow up practice. The DDL exercises were done on a bilingual concordancer using newspaper corpus. It was demonstrated that each DDL approach can be effective for improving grammar basics such as understanding and producing noun phrases. Pre- and post-tests showed students made significant gains using all three approaches, and there does not appear to be any significant difference in effectiveness among the three approaches.

### **Literature Review**

#### **The Use of Corpora to Develop Resources**

The growing speed and capacity of computers have had an impact on language learning, just as they have impacted so many other aspects of modern life. Using large databases of authentic language (called *corpora*, or in the singular, a *corpus*), it is possible to analyze spoken and

written texts in various ways, and this has changed how people understand, teach, and learn language. For example, using a computer and representative corpus, it is possible to determine the most common word in a language, the top 100 or 10,000 most common words, which words appear often together, or what grammatical features are prevalent in particular registers (such as casual conversation or academic writing). In addition to research into language, corpora can be used in resource development. The Collins COBUILD Dictionary, for example, is based on a 2.5 billion-word corpus from globally sourced newspapers, magazines, websites, radio programs and conversations. It provides contextualized definitions so learners can see the definitions, as well as how words are used, what words are often used together, and how frequently these words are used (see “Collins COBUILD,” 2012). Another resource, the *Longman Grammar of Spoken and Written English*, is based on the analysis of a 40 million-word corpus and provides an in-depth description of grammar and how it is used in four registers: academic prose, conversation, fiction, and news reporting (Biber, Johansson, Leech, Conrad, & Finegan, 1999).

Additionally, corpora can be used to teach translation and for second language acquisition. Using a software tool called a *concordancer*, it is possible to explore a corpus of authentic language to observe patterns and make inductive hypotheses about grammatical rules and usage. This is student-centered, discovery-based learning. It contrasts with a traditional L2 classroom setting, which is teacher-centered and where the lessons are explicit (explained) rather than implicit (discovered) and the language data used is limited to whatever is available in the provided textbook or audio / video tape.

### **Computer-Based Versus Paper-Based Versus Combined Data-Driven Learning (DDL)**

Generally, DDL refers to computer-based corpus exercises. A computer-based approach means the student is directly interfacing with a computer, corpus, and concordancing tool. It is student-centered and student-controlled, the data is limitless, and learning is inductive with the student as researcher (Gavioli, 2001). A paper-based approach means students are looking at printed concordance lines on paper, not on computer screens. In other words, someone (the teacher) has already done the searches and is providing the resulting concordance lines for students to examine. The tasks are more readily understood and time-efficient, and do not require student-accessed computers (Boulton, 2010). While there have been many studies looking at the distinct advantages of each, there are few studies comparing computer- and paper-based approaches (Boulton, 2010) and no known studies investigating a combined approach.

### **DDL for EFL Beginners**

Using DDL or corpus-based exercises in the second-language classroom is not new (Aston, 2001; Hunston, 2002; O’Keefe, McCarthy, & Carter, 2007; Sinclair, 2004); however, although the potential for L2 applications using corpora is widely recognized, empirical studies investigating classroom benefits are still limited (Johansson, 2009). According to Boulton (2008a), only 4 out of 50 empirical DDL studies surveyed were conducted with beginner or low-proficiency-level students; he further suggested that “current research encourages the belief that DDL is only useful for advanced learners in a computer laboratory” (Boulton, 2008b). There are several reasons for this. First, concordance lines in only the target language can overwhelm beginner-level students (Tono, 2003). Second, with literally millions of words, it is difficult for learners to know where to begin and what to look for (Tribble, 1997). Third, because many languages have various exceptions to grammatical rules, understanding concordance lines and finding patterns can be challenging (Oghigian & Chujo, 2010). Finally, some students (and teachers) have difficulty using the hardware and / or software (Boulton,

2010) and in some regions, computers for classroom use are not available, and free, easy-to-use concordance tools are not readily available (Chujo & Oghigian, 2008).

To address these issues and investigate the use of DDL with less advanced learners, Chujo and Oghigian (2008) developed and tested a grammar-based syllabus: a series of guided worksheets for pairwork-based tasks with a parallel Japanese-English corpus, and a four-step teaching procedure. Using parallel corpora, they demonstrated that DDL can be effectively used with beginner-level classes. The first language (in this case, Japanese) in the parallel corpora permits comprehension of the target-language concordance lines at a glance; students can therefore focus on grammatical structures. By using a series of specific tasks in a worksheet, students can be directed to a particular grammar pattern and, working in pairs, can support each other. Having the teacher provide an explicit explanation of the grammar pattern after students have analyzed the concordance lines allows them to check their hypotheses and reinforces the powerful discovery aspect of learning. Edited paper-based concordance lines can be used to streamline the data, remove confusing exceptions, and address the issues of a lack of or difficulty with direct use of a computer (Chujo & Oghigian, 2008). Regarding the availability of concordance tools, in this study, a new, free, web-based bilingual concordancer, WebParaNews, was used (previously released as AntWebConc-Bilingual) (Anthony, Chujo & Oghigian, 2011). This is described in more detail below.

### **A Case Study Using Three DDL Approaches in the EFL Classroom**

#### **Computer-Based, Paper-Based and Combined DDL for Beginner-Level Students**

The purpose of this study is to report on the results of computer-based DDL, paper-based DDL, and a combination of both computer- and paper-based DDL for beginner-level university students learning grammar basics. This study is built on the successes of earlier case studies and looks specifically at comparing computer-based, paper-based, and combined concordancing. The efficacy of the syllabus, materials, and teaching procedure has been established in previous studies; thus, the focus of this study is to determine which approach is most effective, and in particular, whether or not having direct contact with a computer is necessary to achieve learning, as measured by student gains in pre- and post-tests. To begin this section, the study is described. Next, the results are presented and then discussed to clarify the learning effect of these approaches. Finally, implications for L2 classrooms are discussed.

#### **Background of the Study**

**Participants.** The participants were three groups of engineering students at a Japanese university in 2009, 2010, and 2011. DDL classes for all groups met weekly for 90 minutes in a CALL (computer-assisted language learning) classroom for one semester. Each group had 22 first-year beginner-level students. The average TOEIC Bridge test score for each group was 69 out of 100 for Group 1 in 2009, 73 out of 100 for Group 2 in 2010, and 67 out of 100 for Group 3 in 2011. These scores are equivalent to a TOEIC score range of 300-350. The majority of students are Japanese males and were 19 years old at the time of participation.

**Corpus.** The corpus chosen for this study was a parallel Japanese-English *Daily Yomiuri* newspaper corpus (Utiyama & Isahara, 2003). Because people or issues in Japan are often referenced, the content in the concordance lines is generally familiar to students. In addition, the non-technical vocabulary featured in this corpus is a target for this level of university study.

**Concordancing Tool.** Until recently, one of the few available tools has been ParaConc, a commercial CD-ROM-based software developed for researchers. This tool was used both with Group 1 in 2009 and by the teacher who created the paper-based DDL for Group 2 in 2010.

(As of 2012, the cost of a two-year, 15-user site license was US\$750 [“Concordancers and corpora,” n.d.].) Although ParaConc proved to be a useful tool, Group 1 students using it in 2009 reported that it took time to adjust the settings for each use, sometimes *kanji* (Japanese characters) in the concordance were garbled, and it could only be accessed in class per the licensing agreement. Additionally, because ParaConc was designed for researchers, the interface takes time to learn and is not necessarily geared toward student use. Thus, for Group 3 students, WebParaNews, a more user-friendly version of AntWebConc-Bilingual including a *Daily Yomiuri* newspaper corpus was used. WebParaNews is built on a standard Model-View-Controller (MVC) framework which works much like a Google search and can be viewed with a standard browser (e.g., Internet Explorer, Firefox, Chrome, Safari, and Opera). There are no initial settings, so searches can be done immediately. It is free, and because it is web-based, can be accessed from home as well as the classroom (WebParaNews, 2012).

### **The DDL Syllabus**

In Japan, students and employees typically take the Test of English for International Communication (TOEIC) to measure English proficiency. The DDL syllabus used in this case study was based on research findings from Chujo (2003), who analyzed the vocabulary in retired TOEIC tests and secondary school textbooks used in Japan and created a vocabulary list addressing the gap. In a later study, Uchibori, Chujo, and Hasegawa (2006) identified grammatical structures found in TOEIC tests and secondary school textbooks. They found that secondary-level texts lacked noun phrases (NPs) and verb phrases, which are prevalent in TOEIC tests. Phrases are an integral element of English; thus, the DDL study was designed to teach NPs.

The DDL syllabus used in this study is shown in Table 1. It started with lexical-based concepts, such as identifying word classes and derivations, and progressed to various types of NPs, e.g., determiner + adjective + noun (*this + business + meeting*) or determiner + noun + to-infinitive (*the + decision + to hire*). Students used the vocabulary taught in each previous vocabulary lesson as the DDL search terms in the subsequent lesson. Thus, vocabulary was always taught first and was spiraled through the curriculum. The vocabulary was also grouped by topic, such as business, personnel, travel, time, and daily life.

**Table 1**  
**DDL Syllabus**

Week	Grammar	Vocabulary
1	(Pre-Test)	Business 1: access, decline . . .
2	Word Classes	Business 2: audit, equipment . . .
3	Derivations & Inflections	Personnel 1: applicant, apply . . .
4	Count & Non-Count Nouns	Personnel 2: accountant, architect . . .
5	NP: adjective (Adj) + noun (N), NP: determiner (Det) + N + prepositional phrase (PP)	Meetings: confirm, convention . . .
6	NP: Det + Adj + N	Travel: arrangement, arrival . . .
7	NP: Det + N + <i>-ing</i>	Money: coverage, credit . . .
8	NP: Det + N + <i>-ed</i>	Buying & Selling: assure, brochure . . .
9	NP: Det + N + <i>to</i> -infinitives	Time: current, due . . .
10	NP: Det + N + <i>who, which, that</i>	Daily Life: appliance, beverage. . .
11	Various NPs	-
12	(Post-Test)	-

### Teaching Procedures

All groups followed the same syllabus and the same four-step teaching procedure. The same teacher (the first author) taught all three groups. Pairs of students explored specific grammatical structures recurring in the corpus. In Step 1, students studied vocabulary with a CALL program. In Step 2, they each had a worksheet with a list of very specific tasks. All groups explored six DDL tasks; however, Group 1 used computer-based DDL, Group 2 used paper-based DDL, and Group 3 used a combination. In Step 3, the teacher provided an explicit explanation of the grammar point or NP, so students could confirm or correct the hypotheses they had made through the inductive DDL tasks. Finally, in Step 4, they practiced the items and consolidated the learning for homework.

### The Three DDL Approaches

**Group 1(2009): Computer-based concordancing.** Students typed a search term into a computer using the ParaConc software, examined the results, and then wrote down an answer on the worksheet. They inductively noticed a pattern and then thought of a hypothesis. A computer screenshot of concordance lines using ParaConc is shown in Figure 1, and a sample inductive DDL task for an NP (Det + Adj + N + PP) is shown in Table 2. In this example, students typed in *a \*opportunity for* (the asterisk indicates a wildcard), looked at the concordance lines, and filled in the missing adjectives for the supplied phrase, which were *good, great, new, and unique*.

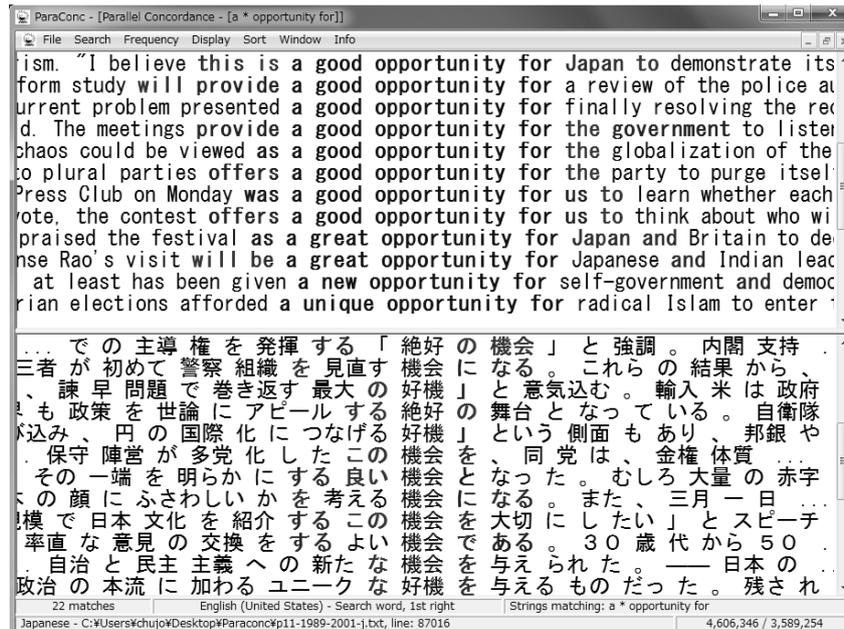


Figure 1. A screen shot for a \* opportunity for using ParaConc.

**Table 2**  
**Sample of a Guided Student Worksheet for Identifying NPs**

Search "a * opportunity for" and find what word comes before opportunity.					
	Determiner	Pre-Modifier	Head Noun	Post-Modifier	
1	a	( )	opportunity	for	Japan
2	a	( )	opportunity	for	a review ...
3	a	( )	opportunity	for	finally resolving ...
4	a	( )	opportunity	for	the government ...
5	a	( )	opportunity	for	the globalization
.	...	...	...	.	...
10	a	( )	opportunity	for	Japanese
11	a	( )	opportunity	for	self-government
12	a	( )	opportunity	for	radical Islam

**Group 2 (2010): Paper-based concordancing.** The Group 2 students used paper-based DDL to follow exactly the same classroom procedure. However, they looked at concordance screenshots printed onto a worksheet, and they did not interact with the computer program. The concordance lines were in color and were from the same parallel Japanese-English newspaper corpus. In the sample task shown in Figure 2, students were first asked to circle the head noun and underline the NP containing the word *lawyer*, including pre- and post-modifiers. In the second task, they looked at various pre-modifier superlative adjectives for the head noun *country* and prepositional phrase post-modifiers, e.g., *in* + place-name. The Japanese translation in the second task was omitted to allow more space for student answers on the handouts.

**Unit 4 Personnel (2) Noun Phrase**

月日      学科      学生番号      &      名前      &

次の 1) から 5) の検索結果を見て、1) の lawyer のように、それぞれの名詞 (head noun とよぶ) を〇で囲み、名詞を説明している (名詞の) 前の部分と (名詞の) 後ろの部分に下線を引いてから書き出そう。

1) lawyer

... revenues. A 67-year-old lawyer has been hired by Aum Super  
 or advisers. A Manhattan lawyer brought a picture to the b  
 ure effective action. The lawyer for the employee's family  
 es or that he served as a lawyer for a drug dealer. At this  
 lf with chloroquine and a lawyer for the plaintiff asked wh  
 nt, I read a book about a lawyer in a wheelchair and it has  
 Law Faculty, and is now a lawyer in Tokyo. His younger brot  
 , the plaintiffs and the lawyer of the defendant will repe

選 弁護 人に 決ま った 大阪 弁護 士 会 の 男性 弁護 士 (67)  
 が、 現在 マンハッタン で 弁護 士 を し て い る ロラ ンド・ト  
 ン 元 社 員 側 の 弁護 士 は 「 組 合 は 沈 黙 し て い  
 麻 薬 取 引 を す る 人 た ち の 弁護 士 だ っ た 一 っ と い っ た 批 判 だ  
 た と み ら れ る 。 原 告 側 弁護 士 「 証 人 も クロ キ ン を 飲  
 こ う 本 で 読 ん だ 車 い す の 弁護 士 に あ こ が れ た 。 自 立 し て  
 た が、 現 在 は 東 京 第 二 弁護 士 会 の 所 属 の 弁護 士 。 三 男 の  
 原 告 側 代 理 人 の 高 橋 勲 弁護 士 の 話 「 双 方 で 時 に は 深 夜

(説明している部分)      +      名詞      +      (説明している部分)

1. <u>A 67-year-old</u>	+      lawyer	なし
2. _____	+      _____	_____
3. _____	+      _____	_____
4. _____	+      _____	_____
5. _____	+      _____	_____
6. _____	+      _____	_____
7. _____	+      _____	_____
8. _____	+      _____	_____

2) country      それぞれの説明部分の品詞構成も考えてみよう。

... age. Singapore is the most affluent country in Southeast Asia, but it a  
 king power, will be the most important country in Europe. That is an endur  
 Japan was considered the most important country in Southeast Asian countri  
 the United States is the most mature country in these four fields and is t  
 urning structure of the most powerful country in the world, and some idea  
 t in 1980, and is the newest independent country in Africa. A neon sign sp

(説明している部分)      +      名詞      +      (説明している部分)

1. <u>the most affluent</u>	+      country	in Southeast Asia
2. _____	+      _____	_____
3. _____	+      _____	_____
4. _____	+      _____	_____
5. _____	+      _____	_____
6. _____	+      _____	_____

Figure 2. Paper-based DDL sample tasks for NPs.

**Group 3 (2011): Combined paper- and computer-based concordancing.** The Group 3 students used a combination of paper-based tasks and computer-based tasks. Instead of the standalone ParaConc, this group used the new online concordancer, WebParaNews. In the exercise shown in Figure 3, students typed in the string *lawyer \*ing* and observed the NP structures for the head noun *lawyer* involving present participles. They filled in the blanks in the exercise, similar to those in Table 2. This group was given two more web-based exercises and three paper-based tasks, such as those shown in Figure 2.

The screenshot shows the WebParaNews interface with the following details:

- Target Language: English (selected), Japanese
- Search term: lawyer \*ing
- Database: news\_corpus\_full
- Sort 1: CENTER, Sort 2: 1L, Sort 3: 1R
- Sampled Hits: 10 of 12

**Target Corpus**

HIT	Target Corpus
<input type="checkbox"/> 1	A lawyer acting as judge is unthinkable in a society
<input type="checkbox"/> 2	Isao Okamura, a lawyer representing an association of crime victims, said, "V
<input type="checkbox"/> 3	A lawyer representing the victim said, "At the time, no
<input type="checkbox"/> 4	n was misleading," said Naoki Fukuchi, a lawyer representing plaintiffs in Tokyo.
<input type="checkbox"/> 5	er for the preservation of its assets, a lawyer serving as the receiver will become responsible for
<input type="checkbox"/> 6	A lawyer working for the plaintiffs also said, "There is
<input type="checkbox"/> 7	Another lawyer defending Takeuchi also denied allegations that Takeuc
<input type="checkbox"/> 8	A Japanese lawyer representing the Philippine woman recently submitted a
<input type="checkbox"/> 9	id they will question a 45-year-old male lawyer belonging to the Yokohama Bar Association who was
<input type="checkbox"/> 10	Yoshio Mitsui, the lawyer representing the victims, said the court's decision to

**Reference Corpus**

HIT	Reference Corpus
<input type="checkbox"/> 1	一方の弁護士が同時に裁判官である事は、法治社会では考えられない事であるからである。
<input type="checkbox"/> 2	犯罪被害者の会代表幹事の岡村勲弁護士は「被害者が悲しんでいる姿を見ないと、加害少年はどれだけ罪が大きい かを自覚できない」と指摘する。
<input type="checkbox"/> 3	弁護士は「当時、助けてくれる人もなく、母親は正常な判断ができる状態ではなかったと思う」と話す。
<input type="checkbox"/> 4	会見に臨んだ東京訴訟の福地直樹弁護士は「副作用と薬害を混同する国民はいないはずで、『薬害根絶』が誤解を招 くという厚生省の理屈は納得できない。
<input type="checkbox"/> 5	山一証券が自己破産を申請し、裁判所から財産の保全命令を受ければ、資産管理や清算業務は破産管財人の弁護士が担 当する。
<input type="checkbox"/> 6	BC総戦犯者訴訟弁護団の平湯真人弁護士も、「行政機関に統一的な窓口はなく、国会議員にも受け皿がない状態。 国民からの働き掛けがないと、動かないだろう」と話す。
<input type="checkbox"/> 7	竹内前知事は「供述調書は検察官の描いた考えを強引に押しつけたもの」とし、弁護人も「わいろで買ったという割 引債は前知事が別の資金で購入した」などと述べ、無罪を主張した。
<input type="checkbox"/> 8	Aさんは弁護士を通じ、フィリピンでの結婚契約書をこのほど都内の区役所に届け出た。
<input type="checkbox"/> 9	示談交渉には、横浜弁護士会の男性弁護士（45）が加わっており、同課はこの弁護士からも事情を聞く。
<input type="checkbox"/> 10	被害者側の三井嘉雄弁護士は「来年四月施行の改正少年法の内容を先取りした取り組み」と評価している。

Figure 3. A screen shot for lawyer \*ing using WebParaNews.

## Data Collection

Students in each of the three groups were given a pre-test at the beginning of the semester and a post-test at the end. There were 60 questions categorized into four groups: identifying an NP using high frequency words (HFW), identifying an NP using more complex TOEIC-level vocabulary (Chujo, 2003), producing an NP with HFW vocabulary, and understanding a TOEIC-type NP. Each question was controlled by word level (a HFW or TOEIC-level word), sentence length, and NP structural pattern (e.g., [Det + Adj + N], [Det + N + PP], and [Det + Adj + N + *to* / *-ing* / *-ed*]). Sample questions showing the four types of questions and answers for NPs are shown in Figure 4. Although the tests employed some of the vocabulary used in the course, the focus was not on vocabulary, but on whether or not students could identify and produce the target grammar, in this case, NPs. The same test was used for both the pre-test and post-test, although the order of the questions was changed to protect assessment validity. Since research has indicated that, due to the time interval between tests and because students are not given the correct answers at any time, using the same test should not impact the end results (Takahashi, Suzuki, & Takefuta, 2003). Using open-ended comment questions, feedback was collected from Group 3 students about the combined DDL, asking which they preferred (computer-based, paper-based, or combined) and why.

1	Identifying NPs (HFW)	<i>Underline all the noun phrases.</i> <u>A famous writer</u> wrote <u>these comic books</u> .
2	Identifying NPs (TOEIC words)	<i>Underline all the noun phrases.</i> We are unable to meet <u>the present demand</u> .
3	Producing NPs	<i>Complete the sentence.</i> My brother studied ( <b>three foreign languages</b> ) at the university. (兄は大学で3つの外国語を勉強した。)
4	TOEIC-type NPs	<i>Choose the best answer.</i> I appreciate your <u>(B)</u> offer. (A) generosity (B) generous (C) generously (D) generousness

Figure 4. Sample pre- and post-test questions.

### Results and Discussion

#### Results from the Pre- and Post-Tests

The results for the NP pre- and post-tests are shown in Table 3. Each test score is shown as a percentage. The score differences between the four types of NP questions were examined and a paired t-test to each pre- and post-test score for the three DDL approaches was applied. There was a significant increase in all types of NP questions in the three varying approaches with a difference significant at the 1% or 5% level. The results clearly show that the students made significant gains with all three DDL approaches. Although the students in 2011 changed from Paraconc to WebParaNews, it is believed that because the syllabus, type of tasks, teaching procedure, corpus, and assessments remained the same, the results are comparable. To confirm this assumption, this variable was measured in a subsequent study (Chujo, Oghigian, & Nishigaki, 2012).

**Table 3**  
*Test Scores (%) by Group for Three DDL Approaches*

Category	Group 1 Computer-based DDL			Group 2 Paper-based DDL			Group 3 Combined DDL		
	Pre-test	Post-test	Gain	Pre-test	Post-test	Gain	Pre-test	Post-test	Gain
Identifying NP (HFW)	53.6	73.3	19.6**	53.6	74.2	20.6**	61.8	76.7	14.9**
Identifying NP (TOEIC words)	43.6	88.5	44.9**	50.3	75.5	25.2**	61.5	88.2	26.7**
Producing NP	57.3	77.0	19.6**	64.9	78.8	13.9**	65.2	74.9	9.7*
TOEIC-type NP	55.5	64.9	9.4*	57.6	66.1	8.5*	53.3	64.6	11.2**

Note. \*  $p < .05$  \*\*  $p < .01$

Next, the pre- and post-test score differences among the three different DDL approaches were examined. To determine if the gains were significantly different from each other, a two-way ANOVA with repeated measures on one factor was performed. The results in Table 4 show that the DDL exercise effect represented by the pre- and post-test scores is highly significant, as in Table 3, with a  $p$ -value less than 0.001 ( $F(1, 63) = 141.1, p < .001$ ). On the other hand, there is no significant difference between the three DDL exercises factor ( $F(2, 63) = 0.905, ns$ ). There

is no significant interaction between DDL approach type and test scores. These results indicate that students showed significant gains with all three DDL approaches, which means all three DDL approaches were effective for teaching the four types of noun phrases shown in Figure 4. However, there is no real difference among the three DDL approaches from which the post-test scores are drawn. This implies that paper-based DDL could be as effective as the other two DDL approaches; therefore, teachers and students can have some confidence that computers are not an essential element for DDL. This is particularly good news for school communities that do not have access to classroom-based computers or computer labs.

**Table 4**  
**ANOVA Summary Results**

Source	Sum of Squares ( <i>SS</i> )	<i>df</i>	Mean Square ( <i>MS</i> )	<i>F</i>
A: Between Groups	400	2	200.0	0.905
Error (S(A))	13914	63	220.9	
B: Within Groups	11504	1	11504.3	141.1****
A x B (interaction)	381	2	190.6	2.34
Error (BS(A))	5136	63	81.5	
Total	31335	131		

Note. \*\*\*\*  $p < .001$

### 2011 Student Feedback

Although similar feedback was collected in 2009 and 2010 from Groups 1 and 2, respectively, feedback from the Group 3 students was of particular interest because they were the only group to use both computer- and paper-based approaches (three paper-based tasks and three web-search tasks in each DDL lesson). In their responses, 32% of Group 3 students indicated they preferred paper-only because they did not like to search by themselves and they liked writing on paper. On the other hand, 32% of students preferred web-search tasks only, because they could also learn about computer searches. Finally, 36% of students preferred having both paper- and web-searches, since they recognized the advantages of both approaches and combining these enabled them to learn effectively, save time, and vary the work. Since roughly one-third of students stated a preference for each approach, it cannot be said that a majority of students showed a strong preference for one over the others; however, it is safe to say that combined DDL, which has the advantages of both approaches, addresses a variety of students' learning styles and preferences.

### Implications for L2 Classrooms

The findings in this study suggest that, with careful guidance (i.e., a teacher-tested worksheet) and supervision, both computer-based and paper-based corpus exercises are equally effective. Clearly, a combination meets the needs of a greater variety of learning styles. Paper-based concordance lines can be edited to remove confusing exceptions, shorten the number of examples, and ensure students are looking at the correct pattern. Paper-based exercises also save a great deal of time, since students are not doing the searches, are not distracted by the computer, do not require hardware or software troubleshooting, and are not unsure about whether or not they are finding the correct data. Alternatively, once students become more proficient with concordancing, they often enjoy using this powerful resource to explore words or structures they are curious about. Asking students to work in pairs has also been beneficial

since they are able to negotiate and agree on hypotheses they are forming based on what they are seeing. At a minimum, pairwork eliminates the single silent student who sits at a blank computer or worksheet not understanding what to do.

These findings also suggest that teachers can use paper-based corpus exercises to teach grammar basics in schools without computers. Although few pre-made corpus-based materials are currently available, it is inevitable that more will be published in the coming years as DDL continues to make inroads in L2 acquisition. Paper-based exercises do require, however, that the teacher or material developer have access to a corpus and concordancing tool to produce them. Until pre-made resources are available, there can be lengthy preparation time involved. This might be minimized by following sample tasks in previously published studies discussed in this paper. For more information on sample tasks, issues to consider, and resources, see Oghigian and Chujo (2010).

The case study presented also supports the contention that a bilingual concordancer is important to use at the beginner level. Low proficiency students tend to be much more relaxed when they have quick and easy access to the translation (Chujo & Oghigian, 2008). They can check the translation and then focus on the structure of the grammar, instead of struggling with the meaning and structure of the concordance line. As noted in the Literature Review, beginner-level students have difficulty coping with monolingual concordance lines. Trying to show them a computer, new software, a new type of task, a new grammar structure, and new vocabulary can be overwhelming for them. For this reason, the bilingual translation helps, as does easing students into the concept of DDL with paper-based concordancing. Once students understand the idea of the task, the transition to a computer tends to be smoother. In fact, as students gradually become accustomed to using DDL and learn to focus just on the grammar structure, they tend to look at the translation less and less.

### **Limitations of the Study**

The participants in this study were engineering students, and it is not known whether this had an impact on their interest or skill in using DDL. Certainly there are numerous studies on L2 DDL (some of which are cited in the Literature Review) in which DDL is used with a variety of first-language students, target languages, and student backgrounds. No studies that look at student majors as a variable of the efficacy of DDL were found in the literature, however. Additionally, researchers in DDL are often the classroom teachers, and it is not known if or to what extent this may impact outcomes.

It should also be noted that one potential problem with using DDL is that students are often more comfortable *studying* about a language rather than *producing* a language. With DDL, the student becomes the researcher and explores the language in a very powerful and motivating way; however, it is essential that production be included so students apply what they are learning to written or spoken constructions.

### **Conclusion**

At the beginner level, DDL has shown to be effective in teaching basic grammar such as NPs, confirming earlier studies (Chujo & Oghigian, 2008; Chujo, Anthony, & Oghigian, 2009; Oghigian & Chujo, 2010). Regarding the comparison of paper-based DDL, computer-based DDL, and combined DDL, the results of this study indicate that there does not appear to be any significant difference in effectiveness among the three approaches. Students showed significant gains with all three approaches, and all three clearly have distinct advantages. Since these inductive corpus-based approaches basically involve the same cognitive processes (e.g.,

noticing, hypothesis formation, and hypothesis testing), teachers can choose the approach best suited for students with regard to class time, learning styles and preferences, and student interest in exploring beyond classroom-defined parameters. In regard to this final point, to date, there has been a need for freeware open-access parallel concordance tools that are designed for use by students and are accessible anytime and anywhere. Such a tool, WebParaNews, has been developed by the authors of this paper and has been tested in the classroom with positive results.

#### **Author Note**

Kiyomi Chujo, College of Industrial Technology, Nihon University, Chiba, Japan; Laurence Anthony, Center for English Language Education in Science and Engineering (CELESE), Waseda University, Tokyo, Japan; Kathryn Oghigian, Center for English Language Education in Science and Engineering (CELESE), Waseda University, Tokyo, Japan; Asako Uchibori, College of Industrial Technology, Nihon University, Chiba, Japan.

This research was funded by a Grant-in-aid for Scientific Research (21320107) from the Japan Society for the Promotion of Science and the Ministry of Education, Culture, Sports, Science and Technology.

Correspondence concerning this article should be addressed to Kiyomi Chujo, College of Industrial Technology, Nihon University, 2-11-1 Shin'ei, Narashino-shi, Chiba 275-8576, Japan. Email: [chujou.kiyomi@nihon-u.ac.jp](mailto:chujou.kiyomi@nihon-u.ac.jp)

## References

- 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, Netherlands and New York, NY: Rodopi Press.
- Aston, G. (2001). *Learning with corpora*. Houston TX: Athelstan.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *The Longman grammar of spoken and written English*. London, England: Pearson Education ESL.
- Boulton, A. (2008a, March). *Evaluating corpus use in language learning: State of play and future directions*. Paper presented at the American Association of Corpus Linguistics, Provo, Utah.
- Boulton, A. (2008b). DDL: Reaching the parts other teaching can't reach? In A. Frankenberg-Garcia, I. Rkibi, M. R. Cruz, R. Carvalho, C. Direito, & D. Santos-Rosa (Eds.), *Proceedings of TaLC 8 – Lisbon, 8<sup>th</sup> Teaching and Language Corpora Conference* (pp. 38-44). Lisbon, Portugal: Associação de Estudos e de Investigação Científica do ISLA-Lisbo Fichier.
- Boulton, A. (2010). Data-driven learning: Taking the computer out of the equation. *Language Learning*, 60(3), 534-572. <http://dx.doi.org/10.1111/j.1467-9922.2010.00566.x>
- Chujo, K. (2003). Eigo shokyuusha-muke TOEIC-go1 to 2 no sentei to sono kouka [Selecting TOEIC vocabulary 1 & 2 for beginning-level students and measuring its effect on a sample TOEIC test]. *Journal of the College of Industrial Technology, Nihon University*, 36, 27-42.
- Chujo, K., Anthony, L., & Oghigian, K. (2009). DDL for the EFL classroom: Effective uses of a Japanese-English parallel corpus and the development of a learner-friendly, online parallel concordancer. In M. Mahlberg, V. González-Díaz, & C. Smith (Eds.), *Proceedings of 5<sup>th</sup> Corpus Linguistics Conference*. <http://ucrel.lancs.ac.uk/publications/cl2009/>
- Chujo, K., & Oghigian, K. (2008). A DDL approach to learning noun and verb phrases in the beginner level EFL classroom. In A. Frankenberg-Garcia, I. Rkibi, M. R. Cruz, R. Carvalho, C. Direito, & D. Santos-Rosa (Eds.), *Proceedings of TaLC 8 – Lisbon, 8<sup>th</sup> Teaching and Language Corpora Conference* (pp. 65-71). Lisbon, Portugal: Associação de Estudos e de Investigação Científica do ISLA-Lisbo Fichier.
- Chujo, K., Oghigian, K., & Nishigaki, C. (2012, November). *Beginner level EFL DDL using a parallel web-based concordancer*. Paper presented at the Pan-Asian Conference on Language Teaching and Learning (PAC) 2012 and 9th FEELTA International Conference, Far Eastern Federal University, Vladivostok, Russia.
- Collins COBUILD dictionaries. (2012). <http://www.collinslanguage.com/collins-elt-learners-of-english/cobuild>
- Concordancers and corpora. (n.d.). <http://www.athel.com/mono.html>
- Gavioli, L. (2001). The learner as researcher: Introducing corpus concordancing in the classroom. In G. Aston (Ed.), *Learning with corpora* (pp. 109-137). Houston: Athelstan.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge, England: Cambridge University Press.
- Johansson, S. (2009). Some thoughts on corpora and second-language acquisition. In K. Aijmer (Ed.), *Corpora and language teaching* (pp. 33-44). Amsterdam, Netherlands: John Benjamins.
- Oghigian, K., & Chujo, K. (2010). An effective way to use corpus exercises to learn grammar basics in English. *Language Education in Asia*, 1(1), 200-214. [http://dx.doi.org/10.5746/LEiA/10/V1/A17/Oghigian\\_Chujo](http://dx.doi.org/10.5746/LEiA/10/V1/A17/Oghigian_Chujo)
- O'Keefe, A., McCarthy, M., & Carter, R., (2007). *From corpus to classroom: Language use and language teaching*. Cambridge, England: Cambridge University Press.

- Sinclair, J. (2004). *How to use corpora in language teaching*. Philadelphia, PA: John Benjamins.
- Takahashi, H., Suzuki, H., & Takefuta, Y. (2003). CALL 教材による自己学習と授業活動を融合させた大学生英語聴解力の養成(<特集>第二言語学習とその支援に関する教育工学研究) [An integration of CALL self-study and classroom activities for improving Japanese college learners' English listening skills (Special Issue: Educational Technology Research on Second Language Learning and its Assistance)]. *Japan Journal of Educational Technology*, 27(3), 305-314. Available at [http://ci.nii.ac.jp/els/110003026504.pdf?id=ART0003487222&type=pdf&lang=en&host=cinii&order\\_no=&ppv\\_type=0&lang\\_sw=&no=1356235559&cp=](http://ci.nii.ac.jp/els/110003026504.pdf?id=ART0003487222&type=pdf&lang=en&host=cinii&order_no=&ppv_type=0&lang_sw=&no=1356235559&cp=)
- Tono, Y. (2003). Corpus wo eigo kyouiku ni ikasu [What corpora can do for language teaching]. *English Corpus Studies*, 10, 249-264.
- Tribble, C. (1997). Improvising corpora for ELT: Quick and dirty ways of developing corpora for language teaching. In B. Lewandowska-Tomasczyk & J. Melia (Eds.), *Proceedings of the First International Conference on Practical Applications in Language Corpora* (pp.106-117). Lodz, Poland: Lodz University Press.
- Uchibori A., Chujo, K., & Hasegawa, S. (2006). Toward better grammar instruction: Bridging the gap between high school textbooks and TOEIC. *The Asian EFL Journal*, 8(2), 228-253.
- Utiyama, M., & Isahara, H. (2003). Reliable measures for aligning Japanese-English news articles and sentences. In E. Hinrichs & D. Roth (Eds.), *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics* (pp. 72-79). Sapporo, Japan: The Association for Computational Linguistics. Available at <http://acl.ldc.upenn.edu/acl2003/main/index.html>