

An Effective Way to Use Corpus Exercises to Learn Grammar Basics in English

Kathryn Oghigian
Waseda University, Tokyo
k_oghigian@aoni.waseda.jp

Kiyomi Chujo
Nihon University, Tokyo
chuuju.kiyomi@nihon-u.ac.jp

Abstract

Computer-based data driven learning (DDL) is becoming gradually more prevalent in L2 classrooms, and studies have shown that DDL exercises can be effective even at the beginner level in teaching vocabulary and grammar (Chujo & Oghigian, 2006, 2008). In addition, paper-based DDL studies are gaining ground (Boulton, 2008b), so that even without computers, students are able to explore concordance lines to identify recurring grammatical features and word derivations and to learn how to induce generalizations from the samples to the language in general. In this study, the researchers have developed a series of computer-based and paper-based DDL exercises focused on grammar basics and using a four-step approach that is both inductive and deductive. Course evaluations indicate statistically significant gains in most areas, and students report an overall positive rating on questionnaires.

The benefits of data driven learning (DDL) or corpus-based exercises in the L2 classroom are widely recognized. Leech tells us that "... the corpus, as an information source, fits in very well with the dominant trend...from *teaching as imparting knowledge* to *teaching as mediated learning*" (1997, p. 2); others point to advantages of using DDL: authenticity; active, student-centered learning where the learner has control of the learning process; and the powerful and motivating discovery aspect in which the learner acts as researcher (Braun, 2008; Huang, 2008; Hunston, 2002; Johns, 1994; O'Keefe, McCarthy, & Carter, 2007; Stevens, 1995). There is no question that corpus-based learning is gaining ground in second language acquisition; however, Boulton's (2008a) survey of 50 DDL studies shows only four at the beginner level. Corpus use at the beginning level, either directly on computers or with paper-based material, is the focus of this study.

In the first section of this paper, several considerations for using DDL in the beginner level L2 classroom are provided. Next, examples of computer-based and paper-based vocabulary and grammar exercises are given in the second section. These exercises have been used in the Nihon University Study, described in the third section. In the final section, the advantages and disadvantages of computer-based and paper-based corpus exercises are discussed.

Considerations for DDL in the Beginner L2 Classroom

The authors posit that DDL can be successful with beginner level students, although there are several important considerations.

Choosing an appropriate corpus. The first and arguably most essential consideration for classroom DDL exercises is choosing the appropriate corpus with regard to size, relevance, and accessibility.

Size. The largest and certainly most accessible corpus is “the Internet,” estimated at 100 trillion words in 2007 by Peter Norvig, Director of Research at Google (Norvig, 2007). The use of an immense corpus this size risks “drowning [learners] in data” (Tribble, 1997, p. 2); this risk is present even with corpora of only a few million words. Comparatively smaller corpora, such as the British National Corpus (BNC) with 100 million words, or the Corpus of Contemporary American English (COCA) with 410 million words, can be somewhat more manageable. With these corpora, it is easier to narrow the focus to a particular and more practical language function, context, or register. For example, Biber (2009) points out that conversation generally contains numerous embedded clauses, whereas writing contains many more phrases, especially noun phrases. Therefore, the focus of a communication or writing class might be directed to smaller and more relevant, specialized spoken or written corpora, respectively. Sub-corpora can further narrow the focus to types of spoken or written discourse, for example, academic discourse (e.g., from the Michigan Corpus of Academic Spoken English (MICASE)), or academic-sourced writing (e.g., the COCA academic register). Tribble (1997) advocates creating micro-corpora, but this may be overly ambitious for many teachers, particularly given copyright issues, and may be unnecessary, given the accessibility of many corpora and corpus tools. It is also important to note that with a corpus that is too small, there may not be a sufficient number of examples for learners to see a pattern. A sample of corpora and corpus tools is shown in Table 1; for a comprehensive list of corpora, corpus tools and related information, see David Lee’s *Bookmarks for Corpus-based Linguists* website.

Table 1
Sample of Corpora and Corpus Tools

Corpus, Tool	Access	Corpus Size	Comment
Corpus.BYU (Brigham Young University)	web-based; free	COCA: 410 million words; BYU-BNC: 100 million words; Oxford: 37 million words	This site provides both a tool and several corpora.
Collins	web-based; commercial product	56 million words	This is a tool and a corpus.
Paraconc	commercial product	tool only; a corpus must be uploaded	This is a parallel corpus tool which can display two languages.
Antconc	free; downloadable from the web	tool only; a corpus must be uploaded	This works for PC, Mac and Linux.
MICASE	web-based; free	1.8 million words	This is a specialized corpus of academic spoken English.

Relevance. The choice of corpus should be one that has some relevance to the learners (Braun, 2008). With regard to relevance, the BNC is primarily adult British English (Nation, 2004), with several sub-corpora in commerce, the arts, science, and other fields, which would provide vocabulary and structures in those specific contexts. It would not be pertinent to, for example, children’s vocabulary. COCA provides results in various registers, such as academic journals,

news, magazines, fiction (including movies), and spoken English. The written language of academic journals tends to be formal and technical; newspapers are less technical than academic writing, but still often use formal language; magazines are often informal and contain a high level of colloquial language; and fiction includes creative writing. Spoken corpora, depending on the source, can contain both formal English and slang. Thus with both the BNC and COCA, it would be possible to narrow searches to particular sub-corpora or registers.

In addition, there are many types of specialized corpora, for example, discourse (Cambridge and Nottingham Corpus of Discourse in English [CANCODE]), academic spoken English (Michigan Corpus of Academic Spoken English [MICASE]), magazines (*TIME*), dictionaries (Oxford), and historical language (Corpus of Historical American English [COHA]; Helsinki Corpus). Furthermore, there are corpora used for translation; parallel corpora, which show two languages; and learner corpora, which are collections of text from learners. (For a discussion on types of corpora, see Hunston, 2002, pp. 14-16.)

Accessibility. Additionally, the choice of corpus should be one that is logistically accessible to learners (i.e., easily accessed on the Internet or through student centers).

Controlling the task. The next consideration is control over the task. Although the term *discovery learning* is generally understood to apply to student-designed and interest-driven “raw corpus” searches in which neither the student nor the teacher knows the outcome (Hunston, 2002), there is still a discovery aspect for the learner who can find recurring patterns in pre-vetted corpus searches. There are inherent disadvantages in turning students loose on a language such as English, with its huge number of variations and exceptions, and the potential for confusion and frustration is high. Teachers risk losing “an aura of expertise” in situations in which students become mired in complex language forms that teachers may not be prepared to address; this is particularly significant in areas such as Asia, where teachers are viewed by students as experts rather than facilitators (Abdon & Raab, 2001). Huang (2008, p. 20) suggests a “guided dynamic partnership” in which teacher and student work together and the learning process is neither wholly teacher controlled nor student controlled.

Bridging the gap between research and production. The third consideration has to do with bridging the gap between student as *language researcher* and student as *language producer*. One of the inherent disadvantages in corpus-based learning is that it tends to underscore the idea of L2 as “the study of” as opposed to “the use of.” Teachers might find it even more difficult to encourage production with students who are more comfortable with an arm’s length analysis of language. For that reason, it is vitally important to have a communicative aspect in which to apply acquired knowledge, whether it is in a spoken or written context. Braun refers to this process as moving from corpus to discourse and suggests the importance of using corpus exercises in conjunction with “comments and exercises, explanations, exploratory tasks, study aids and didactic hints for learners and teachers” (Braun, 2008, p. 55).

One final concern involves instructional limitations: Not all programs have the IT infrastructure to support classroom DDL. In schools without computer labs or CALL classrooms, the teacher can use one laptop connected to an LCD projector to demonstrate how to use corpus tools, and then ask students to try searches at home or on library computers. Alternatively, another option is to set up one laptop in the classroom and have small groups or pairs do searches while other groups are engaged in other activities. If no classroom technology is available, paper-based exercises can be used. In this paper, several corpus exercises that could be presented as paper-based in classrooms without IT access are offered.

Using DDL Tasks at the Beginner Level

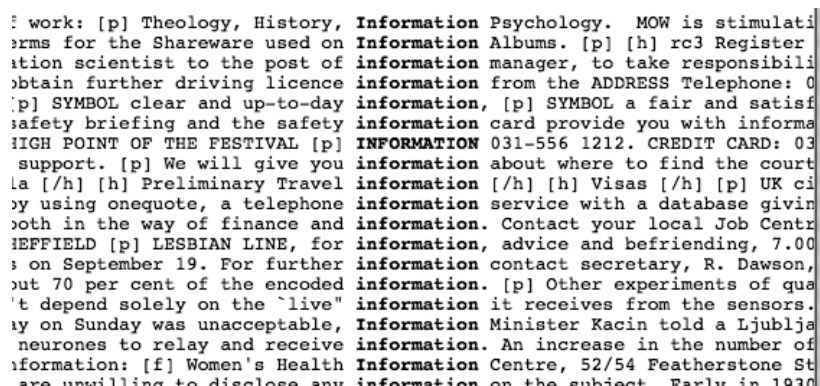
There are many computer-based and paper-based vocabulary and grammar related tasks that can be used successfully at the beginner level. An important point to note is that concordance lines, particularly monolingual concordance lines in the target language, can be overwhelming to beginner level learners (Tono, 2003). For this reason, most tasks at the beginner level are focused on just observing the key word in context (KWIC), or within a context of one to three words on either side of the KWIC. Generally, a beginner level learner is not asked to read or understand the concordance line.

In the first example, shown in Figure 1, students can search the stem of a word with a wildcard (*), for example, *develop**, and can easily see derivations and inflections appearing as a highlighted and aligned KWIC in the center of the data. This is shown using the Collins concordancer. The purpose of this task at the beginner level is to observe the various forms of the word *develop*: *developed*, *developer*, *developing*, *development*, *developmental*, *developments*. This type of task would be useful for understanding derivations of assigned vocabulary words, identifying parts of speech for each form, or to verify the correct form or spelling of a new vocabulary word. This task could be made more advanced by having students also study the use of each form in context; however, it may be advisable to create a paper-based version in which all of the more complex strings could be removed. Generally, however, DDL tasks at the beginner level, particularly for first-time users, may be more successful if the focus remains on only the KWIC, or from one to three words on either side of the KWIC, and not a larger context.

Traditions change and **develop** and so does the church. It us
for short periods to **develop** new ideas. Visitors can see th
hen "go it alone", or **develop** alternative links with other
s the issues and then **develop** a strategy," he says. A centr
ntial nurturer, might **develop** so quickly that she will lose
] A triangles game to **develop** and consolidate the idea of f
h time to erm explore **Develop** new skills Cos that's what it
w black gospel music **developed** away from staid Christian wor
well routine in the **developed** world. If problems arise, do
the start of Lent he **developed** the idea, claiming that "cel:
t when he was 11 and **developed** a wide range of sporting inte
ess. But I have also **developed** a method that may be of some
engineers designed and **developed** an amp and speaker combinati
preneur hotelier and **developer** Sol Kerzner. The main restaur
se include: [p] [p] **developing** and testing medicines for hur
dly accused Iraq of **developing** nuclear weapons technology --
rom the rear. It is **developing** a system that would allow joi
to this meeting for **developing** countries. In the same year,
art did you play in **developing** the [ZF1] thi [ZF0] this cou
etics of Drosophila **development**. [p] [p] 3 Membrane Chemist
development world. This is an exciting
e greatest hope-the **development** of the indoor riding school
central role in the **development** of Europe. I believe all the
Procrustean bed of **development** economics, as, for instance,
zation in a unified **development** strategy", in Meier, Leading
he result of uneven **development**, rather than its cause, the
es of services. The **development** of the "problem" of child al
ey field. Jobs as a **development** officer in local schools and
rits or harm of the **development** having proceeded in the rec
must admit that the **development** of foreign links is importa
ciple was the "full **development** of each individual and the
study of perceptual **development** continues to be the nativis
e end of January, a **development** that has been aggravated by
ment the industrial **development** and so on and the decline o
anning and Economic **Development** Committee combines the func
hether it came from **development** a separate bid completely o
sible, the initial **developmental** trajectory within the real
h) [p] Educational **developments** in England and Wales : Prim
t all the exciting **developments** that had flowed from Medell
ill handle all the **developments** of our products and approva.

Figure 1. Task: Search *develop** and list all the forms of this word that you find. What is the part of speech for each word?

A second fairly easy task for beginner level students would be to search for relevant nouns (i.e., new vocabulary or those noted with errors in student production) using the word plus a wildcard, for example, *information** to understand countable and uncountable nouns. This is shown, again with the Collins concordancer, in Figure 2. If assigning nouns that also function as verbs (e.g., *research*, *program*), it is important to point out to students that they are looking only for plural forms of nouns. In Figure 2, it is very clear that the noun *information* is an uncountable noun, just as a search for *passenger** would clearly show numerous occurrences of *passengers*.



The screenshot shows a list of search results for the word 'information'. The results are displayed in a two-column format. The left column contains the context (a sentence or phrase) and the right column contains the word 'information' followed by its frequency in parentheses. The results are as follows:

Context	Frequency
work: [p] Theology, History, Information Psychology. MOW is stimulati	1
irms for the Shareware used on Information Albums. [p] [h] rc3 Register	1
ation scientist to the post of information manager, to take responsibili	1
obtain further driving licence information from the ADDRESS Telephone: 0	1
[p] SYMBOL clear and up-to-day information , [p] SYMBOL a fair and satisf	1
safety briefing and the safety information card provide you with informa	1
HIGH POINT OF THE FESTIVAL [p] INFORMATION 031-556 1212. CREDIT CARD: 03	1
support. [p] We will give you information about where to find the court	1
la [/h] [h] Preliminary Travel information [/h] [h] Visas [/h] [p] UK ci	1
oy using onequote, a telephone information service with a database givin	1
oth in the way of finance and information . Contact your local Job Centr	1
IEFFIELD [p] LESBIAN LINE, for information , advice and befriending, 7.00	1
s on September 19. For further information contact secretary, R. Dawson,	1
ut 70 per cent of the encoded information . [p] Other experiments of qua	1
't depend solely on the "live" information it receives from the sensors.	1
y on Sunday was unacceptable, Information Minister Kacin told a Ljublj	1
neurons to relay and receive information . An increase in the number of	1
information: [f] Women's Health Information Centre, 52/54 Featherstone St	1
are unwilling to disclose any information on the subject. Early in 1930	1

Figure 2. Task: Search nouns such as *information**, *passenger**, *homework**, *data**, *research**, *language**, and *program**. Which of these nouns are countable? Which are not countable? (Only *information** is shown due to space limitations.)

Beginner level students can also look for patterns with modifiers (adjectives and adverbs), and phrases (noun, verb, and prepositional). In all of these cases, the tasks can be done as computer- or paper-based exercises.

The next tasks use the BYU COCA corpus. Each corpus tool provides query syntax used in searches. Basic syntax used for the BYU COCA corpus include *[at*]* to search for articles, *[ii*]* for prepositions, *[n*]* for nouns, *[v*]* for verbs, *[jj*]* for adjectives, *[rr*]* for adverbs, and *[rrq*]* for *wh*-clauses.

In Figure 3, a search of the BYU COCA corpus for *a * organization* shows that an adjective comes between the article and the noun, as in a *nonprofit organization*, a *national organization*, or a *new organization*. For a more advanced task, students can be asked to write a sentence using a noun phrase (NP) with the head noun *organization*.

CONTEXT	
A NONPROFIT ORGANIZATION	
A TERRORIST ORGANIZATION	
A NATIONAL ORGANIZATION	
A NON-PROFIT ORGANIZATION	
A PRIVATE ORGANIZATION	
A NEW ORGANIZATION	
A RELIGIOUS ORGANIZATION	
A NEWS ORGANIZATION	
TEXT (KWIC)	[FREE] FREQUENCY LIST
CONTEXT	
	of his son with girlfriend Anna Hansen, 28, who works for a nonprofit organization . " Wassup, world? " the cyclist wrote on beha
ntury	School of Music. He founded and serves as the artistic director of a nonprofit organization , Soli Deo Gloria, which sponsors the c
ntury	ministers at the conference in Florida. " The idea you would run a nonprofit organization without the best people is crazy! " Wel
olitan	, " says Helga Luest, president and CEO of Witness Justice, a nonprofit organization that helps survivors of violence and trauma
iment	and his actions. He globe-trots, rescues drowning kids, and runs a nonprofit organization with Law &; Order's Jesse L. Martin (v
	from Tanzanian cheesemakers to Romanian admen. Somehow she found time to co-found a nonprofit organization devoted to H
3azaar	" Still vexed? Check with the Authentics Foundation (myauthentics.com), a nonprofit organization dedicated to helping consume
nian	care has convinced thousands of Americans to deal with their own dead. A nonprofit organization called Crossings (www.crossir
intry	former government-policy adviser and, since 2000, the president of Riverkeeper, a nonprofit organization that works to safegua

Figure 3. Task: Search *a *organization*. What type of word often appears between the article *a* and the noun *organization*?

In Figure 4, a search for *[at*] clear [n*]*, shows that the adjective *clear* appears in a basic noun phrase (article + adj. + noun), such as *a clear message*, *a clear view*, or *a clear understanding*. A more advanced exercise might be to ask students to produce a sentence for each NP using *clear*. It is notable that the BYU corpus tool also gives these phrases in order of frequency of occurrence in this 410 million word COCA corpus, so students might expect to see these particular phrases with the adjective *clear* more often than other phrases.

A CLEAR MESSAGE	248
A CLEAR VIEW	227
A CLEAR UNDERSTANDING	209
A CLEAR PICTURE	190
A CLEAR DAY	182
A CLEAR SENSE	171
A CLEAR SIGNAL	134
A CLEAR IDEA	133
[FREE] FREQUENCY LISTS	
More information...	
the Cubans. MARTIN: OK. FORERO: So, that's a clear message to the Obama administration. MARTIN: Jesus, let's have the	
calling the shots here. In my view, he was trying to send a clear message , first, to chill the Obama diplomatic initiative, not to kill it	
in to the club. " Her slight emphasis on the last two words was a clear message -- if he was going to be a no-show, far be it from	
" The child's wide eyes sparkled. # Stillshadow raised her palm in a clear message : By remaining silent. She scrawled in the dirt, muttering to herself	
signing this bill today," Obama said, " I intend to send a clear message : that making our economy work means making sure it works for everyone;	
meant being late for a soccer game or dance rehearsal, but it sent a clear message that worshipping together is a priority. " To keep from portraying h	
day before. No one was killed by the rockets, but they sent a clear message that Israel's offensive has failed to extinguish Hamas's ability to launch e	
for Yisrael Beytenu. It was refreshing, with an attitude, and with a clear message , " said Israeli pollster Yitzhak Katz. " There was an atmosphere of	
Qaeda is trying to put down roots. The president himself should also send a clear message to the Muslim world that the United States is at war not w	
made headlines. The volume of people hospitalized and the number of deaths sent a clear message that polluted air could no longer be ignored. " Th	
to use practice time wisely, with my practice chart, I was sending a clear message that the only aspect of practicing that counts is time served. # I	

Figure 4. Task: Search *[at*] clear [n*]*. What kind of phrase is this?

In Figure 5, a search of the verb *enjoy [v*]* shows that it is followed by a gerund.

CONTEXT	
	ENJOYED BEING
	ENJOYED WORKING
	ENJOYED WATCHING
	ENJOYED READING
	ENJOYED PLAYING
	ENJOYED DOING
	ENJOYED HAVING
	ENJOYED TALKING
[FREE] FREQUENCY LISTS	
	high. When she woke up would do bong hits. It seems like she enjoyed being high for a long time and suddenly it no longer
	, and she woke up, would do bong hits. It seems like she enjoyed being high for a long time, and then suddenly it was no lo
	all of the above. # Under any number of circumstances, Kimberly van Dorn enjoyed being the center of attention. Last night
sep	. # Right now her job was the main focus in her life. She enjoyed being the Executive Publicity Director for Eclectic Records.
	twisted ankle. She mourned the woman she had been. How could she have enjoyed being the slave of that microscopic tyra
	to scare people with their letters, but then again, Paul could not have enjoyed being told that his physical presence was wea
	and we believe that the economy will eventually turn around. " Though she's enjoyed being a stay-at-home mom to their so
	left them with a feeling I liked -- being stuffed to the gills. I enjoyed being with my mother, but in those years my bid to be
	a job that she loved. She loved working with young people. She really enjoyed being a teacher. She seemed blissfully happy
	. Every, every single Saturday I volunteered to help him out just because I enjoyed being around him. : (Voiceover) To his p
	in his footsteps. : I wanted to be a businessman. so I enjoyed beina involved in all of that. ! 5:5755: : (Voiceover

Figure 5. Task: Search *enjoyed* [v*]. What kind of verb phrase follows it (gerund or infinitive)?

In Figure 6, a search of another verb, *discuss* [rrq*], shows the *wh*-clauses that follow it. As more advanced tasks, students can create sentences, stories or dialogues using these language forms.

CONTEXT	
	DISCUSS HOW
	DISCUSS WHY
	DISCUSS WHERE
	DISCUSS WHEN
	TOTAL
[FREE] FREQ	
	'as doomed from the beginning perhaps, all of that we're going to discuss when we come back right after this. Sou
	fears, and beliefs. In addition, families often need a chance to discuss when , where, how much, with whom, and w
	nited States Army and I spend a year in Baghdad. Im just calling to discuss when I was in Baghdad, I found a very
	is behavior. However, he would not address specific allegations in the report or discuss when he stopped taking perf
	re in the United States. # A Chase representative said the bank would not discuss when the cards will be available t
	iter, Muehr saw Macht for the results from the lymph nodes tests and to discuss when he could return to the Tour. T
	very reliable method. Many e-commerce sites and other businesses would rather not discuss when they're attackec
	And my views on these issues I will be more than happy to discuss when the Senate has my confirmation hearing. I
	ad happened there, something that the older inhabitants of the village still liked to discuss when topics for gossip w

Figure 6. Task: Search *discuss* [rrq*]. What *wh*-clauses appear after this verb?

DDL Case Study

The Nihon University Study

The tasks given in this paper have been used successfully in the Nihon University Study, which has been ongoing since 2005. The impetus of this study dates back to 2003, when Chujo (2003) compared the vocabulary in the Test of English for International Communication (TOEIC) with vocabulary taught in junior and senior high school (JSH) textbooks widely used in Japan, and created a vocabulary list to address the gap. Similarly, in 2004, Uchibori, Chujo,

and Hasegawa (published 2006) identified grammatical structures appearing in retired TOEIC tests, and compared these with grammatical structures taught in JSH textbooks. The vocabulary and grammar became the basis for the Nihon University Study. The most notable finding from the Uchibori et al. study was the prevalence of noun and verb phrases appearing in TOEIC, but absent from the JSH texts. As noted earlier, phrases, particularly noun phrases, occur frequently in written discourse (Biber, 2009). Because phrases are an intrinsic element of English, the Nihon University Study was set up to teach noun and verb phrases, derivations, and countable/uncountable nouns using DDL, and to assess student knowledge through pre- and post-test scores. In addition, feedback was also collected from students regarding the usefulness and general impressions of the DDL tasks.

In the first year of the study, part of the vocabulary was taught in a four-week DDL program with two classes of weekly 90-minute lessons of beginner level EFL engineering students. Based on enthusiastic feedback from students, DDL-based grammar exercises were added in 2006 and the program was expanded to ten weeks (a total of 15 hours). In 2007, the study was expanded to two semesters for a total of twenty weeks (30 hours). Students continued to respond favourably and showed gains between pre- and post-test scores (see Chujo & Oghigian, 2006, 2008). The study has been run each year, with subsequent pedagogical modifications.

The 2008 Study

Although student gains in pre- and post-test scores have been noted, the researchers were curious about how the assessments of the DDL students would compare to those of a more typical, traditional English class. In other words, the DDL students could now identify and produce noun and verb phrases, among other tasks, but would students in a typical non-DDL class be able to do the same? In 2008 and 2009, the researchers collected and analyzed assessments from both types of groups. Because pre- and post-test data for 2009 is still being analyzed at the time of this writing, the data for the 2008 study is given in this paper. However, 2009 student comments regarding a paper-based pilot study will be introduced and discussed in a later section.

Participants. The participants were two groups of beginner level engineering students attending the same Japanese university. The targeted students for the DDL class were 22 first year Japanese university engineering students with an English proficiency ranging from TOEIC 290 to 520. The non-DDL class was comprised of 25 first year engineering students, with slightly lower level TOEIC scores. Both groups attended one 90-minute class each week, for a total of twenty weeks over two semesters. The majority of students were male, Japanese, aged 19. The groups were created through the university registration system, that is, they each represented a typical class of students at the beginner level in the English program, but students in the non-DDL class used a listening-focused CD-ROM program (called *First Listening*) and the other studied basic grammar using DDL exercises. Both classes studied the same vocabulary material using a CALL program.

Tasks. The students in the DDL class worked in pairs to explore targeted recurring grammatical structures in a bilingual newspaper corpus (Utiyama & Isahara, 2003), using a parallel concordancer (Paraconc, 2004). Following guidelines on a worksheet, they searched for and examined basic grammatical structures, formed hypotheses about the structures, discussed these with their partners, and recorded their findings. Next, the teacher provided an explicit explanation so students could confirm or correct their hypotheses. For consolidation, practice and production, students then completed follow-up exercises. The syllabus included word classes and derivations and various noun and verb phrases. In addition to the types of exercises

shown in Figures 1 through 6, the task in Figure 7 illustrates the parallel corpus being used to identify adverb location.

Computer-based exercises were used for all 2008 lessons. The tasks were clearly defined, narrowly focused and written in instructions that were followed by pairs of students. This type of reciprocal learning required them to negotiate, and share and discuss their findings. As an additional benefit, any student who became lost had a partner for guidance, and was less likely to sit passively in mute confusion.

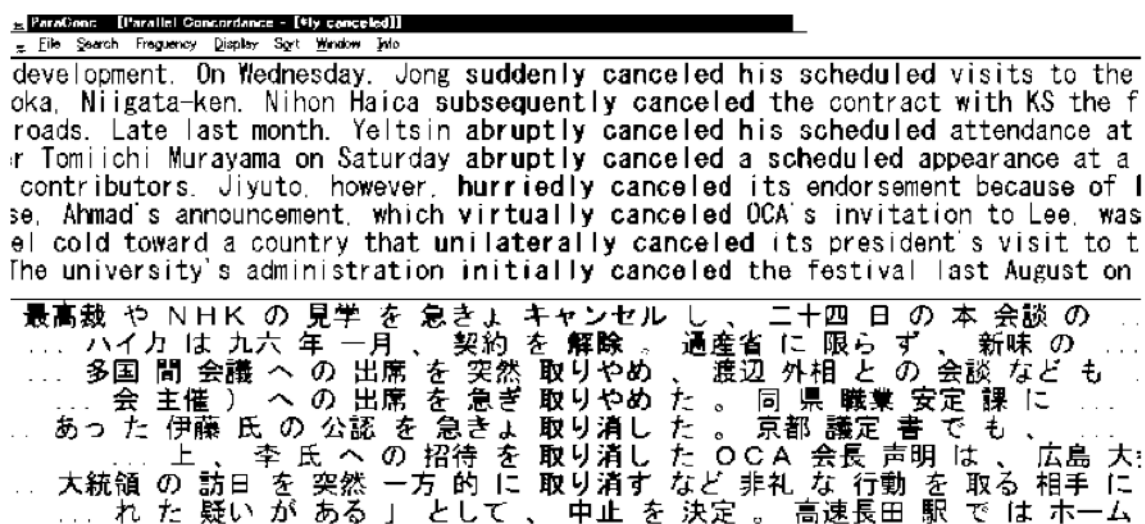


Figure 7. A computer-based task focused on adverb location.

Corpus. With regard to the considerations discussed earlier, the corpus chosen for this program was a Japanese-English bilingual *Daily Yomiuri* newspaper corpus. Students are generally familiar with the ideas appearing in the concordance lines since they often refer to Japanese people or issues. This corpus is also relevant in that it features the type of general non-technical vocabulary targeted at this university level. Using a parallel corpus tool, students can glance at the Japanese lines to get a general understanding of the meaning of the English concordance lines, and then can focus on the *structure*, rather than only the *meaning*, of the target grammar.

Learner Feedback. Learner feedback on the DDL course in general, on the DDL worksheets, grammar explanations, follow-up activities, and teacher feedback was generally positive. Students were asked to give responses on a five point rating scale using "Strongly Agree" (rating 5), "Agree" (rating 4), "Neutral" (rating 3), "Disagree" (rating 2) and "Strongly Disagree" (rating 1). To simplify the results, "Strongly Agree" and "Agree" were combined in Table 2 below, as were "Disagree" and "Strongly Disagree." The results in Table 2 show that most students found the DDL course to be novel, appealing and useful, and found value in the related activities.

Table 2
Learner Feedback, 2008

Activities	Evaluation	Agree	Neutral	Disagree
DDL course	It was novel.	74%	21%	5%
	It was appealing (it was easy to maintain concentration).	68%	26%	5%
	It was useful for grammar learning.	77%	14%	9%
	It was useful for vocabulary learning.	73%	18%	9%
DDL worksheets	They were useful.	53%	32%	16%
Grammar explanations	They were useful.	74%	26%	0%
Follow-up activities	They were useful.	84%	16%	0%
Feedback from teacher	It was useful.	84%	16%	0%

To assess the effectiveness of the program, pre- and post-tests were given to both groups (Figure 8). Students were asked six types of questions regarding vocabulary, word classes, derivations, identifying noun and verb phrases, and understanding more complex noun and verb phrases often found on TOEIC tests.

① Provide the POS (part of speech) for each word.

1. competitive 形容詞 (adjective) 3. requirement 名詞 (noun)

2. the 冠詞 (article) 4. always 副詞 (adverb)

② Write the plural forms of the following words.
If they are uncountable nouns, make a ✓.

1. apple apples 3. baggage ✓

2. family families 4. equipment ✓

③ Circle the head nouns and underline all the noun phrases.

1. The secretary ordered some paper.

2. The agent has forwarded the discount tickets to his new address.

④ Choose the best answer and circle it.

1. During the winter many people enjoy (play / to play/ playing) a variety of indoor sports.

2. Thailand proposed (that / if / what) it would host the meeting.

Figure 8. A sample of pre- and post-test questions.

The gains (i.e., only the increases between the pre- and post-tests) are shown in Table 3. Students in both groups made significant gains at the 1% level in vocabulary, word classes and derivations. The non-DDL students were not able to identify noun or verb phrases, or understand more complex TOEIC-type questions. Students in the DDL group showed gains for all tasks.

Table 3
Student Gains for 2008

Types of questions	DDL + Vocabulary Group	Listening + Vocabulary Group
(1) Vocabulary	33**	16**
(2) Word classes	20**	14**
(3) Derivations	13**	13**
(4) Structure of NPs	17**	3
(5) Structure of VPs	11**	7
(6) TOEIC-type questions	17**	1

***denotes a significant gain at the 1% level*

Computer-Based vs. Paper-Based Exercises

Educators have noted many advantages to computer-based DDL: The approach is student-centered and student controlled; the learning is inductive with student as researcher; there is a limitless supply of data; and corpus-based learning strategies can be applied elsewhere by students (Leech, 1997; Stevens, 1995; Tomlinson, 1998). However, not all schools have facilities for computer-based learning, and not all teachers and students are comfortable with technology-based instruction. An alternative is paper-based corpus exercises. By removing computers from the interaction, it is also possible to remove the additional burden on the student of grappling with computer use *and* the software manipulation *and* the task at hand (Boulton, 2008b) as well as other potential student- or teacher-technophobic issues. Paper-based work can also provide a means to ease learners into corpus-based learning (Boulton, 2008b).

Another important advantage of paper-based corpus work is that it is also easier to control the focus by “filtering” or editing concordance lines (Braun, 2008; Huang, 2008; Hunston, 2002). The use of printed results helps ensure that all students in the class, particularly novice students, are literally “on the same page.” For example, if the goal is asking beginner level students to search for an adjective, such as *clear*, to identify a simple noun phrase [article + adjective + noun], edited paper-based concordance lines can eliminate occurrences of *clear* used in other structures such as verb forms, or before a relative clause. Without editing, the teacher would need to be prepared to discuss various usages of this word, including, for example, the phrase *steer clear*, as shown in Line 21 of Figure 9.

find any isolated and clear teaching about the kind of
 the windscreen crystal clear. They stay supple in the
 gripes and hopes. So clear the confusion with our ex
 ing networks and state clear policies for their protec
 laffe which, although clear, is still tentative. [p]
 sky Nice day he said Clear. Sniper weather. A good d
 : it does amount to a clear statement of Latvian sove
 il Gorbachov. It is a clear sign that he has won the
 Kinnock, has made it clear that the overwhelming vot
 ential election, it's clear there will have to be a r
 die says they made it clear there were conditions: [h
 yond the barbed wire, clear in his binoculars, Adam o
 e testimony, with its clear implication that all bad
 so win" or "I want to clear hellip; feet." The athlet
 edict and a legal all-clear, the police advised the p
 igh interest loan to clear arrears. REMEMBER - TALK
 a mantle of her snow, clear, astonishing, pure - the
 leader Robert Dole as clear favourite to win the Repu
 clamping companies to clear their land. [p] [h] Brit
 d, leaving him with a clear run on goal. That miss, o
 any's products. Steer clear. Only an independent fina
 League giving them a clear advantage over Cambridge,
 I had this wonderful clear road to a writing career.
 NE Train crash men in clear SECTION Home news STORY B
 fixed, but it is not clear what form it will take. O
 ne fans by refusing a clear 63rd-minute penalty claim
 remier League made it clear at the meeting yesterday
 age of hindsight, the clear physical evidence of caus
 1 them, though it was clear he was annoyed by it. But
 1991--dawned warm and clear. Morning light seeped ac
 dictory emotion. It's clear she wants attention but i
 e. At any rate, it is clear that the qualities attrib
 Jennings Bryan, was a clear violation of the spirit,
 of Kuwait. But it's clear that the Kremlin is now p
 becoming increasingly clear that this is the end of t

Figure 9. Unedited concordance lines for *clear*.

Nihon University Study 2009 Student Feedback

As a pilot program in the Nihon University Study for 2009, part of five lessons in the second term were given as paper-based tasks, and student feedback in the form of anonymous comments was collected. In these five lessons, students looked at only printed monolingual concordance lines from a Japanese-English newspaper parallel corpus. Tasks were given both in colour and in black and white. A sample exercise is shown in Figure 10.

afraid を含む動詞句の部分に下線を引き、書き出そう。

we need not be afraid of such an heinous challenge, and not to be afraid of the pain." An income tax o
 market. Are they afraid of resuscitating the memories
 It might be too afraid of street protests by unions.
 ent years. I am afraid that in the eyes of Western b
 economy. We are afraid that may impair the economic
 that they were afraid that if they made a complaint
 mer and we were afraid the Soviet Union would try to

※ *afraid* の後に共通して現れている語は何でしょう。

**able* と **ful* を含む動詞句の部分に下線を引き、書き出そう。

economic recovery are visible in those countries
 in an election. Art is useful in restoring a des
 ustrialized nations is capable of giving the nece
 all Government." We are doubtful of the agency's
 : goods. "His works are wonderful since they are
 ently taxable, some are doubtful that the reform
 al exhaust emissions are harmful to the health o
 recent increase is attributable to the present

be afraid of... _____

be afraid of the pain _____

Are... afraid of ... _____

be too afraid of street ... _____

am afraid that... _____

are afraid that ... _____

were afraid that _____

were afraid (that) _____

of, that 形容詞の補語

are visible in _____

is useful in ... _____

is capable of ... _____

are doubtful of ... _____

are wonderful _____

are doubtful that _____

are harmful to ... _____

is attributable to ... _____

Figure 10. A paper-based exercise focused on adjectives.

With regard to the advantages of using paper-based material, students reported that using paper-based material saved a lot of time and more tasks could be completed during class, that it was easier to understand the grammar and they could do so more quickly (because the concordance results were presented to them), and that they did not have to worry about finding a different or wrong outcome since the handouts provided the “correct results.” Students also found it easier and more immediate to write directly on [more familiar] paper, so this made the task easier. They liked that they could review or re-do answers at home, and generally felt that fewer mistakes made for a smoother class.

Comments regarding a preference for computer-based work included the ideas that the work was active, searching the corpus themselves was very powerful, and the learning was “fixed in their memories.” Students preferring computer-based work appreciated having control over the learning process and could therefore explore additional items or interests. They liked having the Japanese translation from the Japanese-English parallel corpus using Paraconc, and disliked that the context was cut shorter than regular concordance lines to fit on the paper. They commented that doing the search gave them time to think about the process and to consider the target vocabulary and grammar more carefully. They appreciated being able to sort left and right, and realized it was easier to memorize spellings when using a computer, because they were physically typing in the words. Some students also complained that underlining and circling on paper-based exercises were monotonous, and for these reasons, they preferred computer-based exercises. Finally, they felt that unless they themselves searched, they would not know how to use corpus tools.

Paper-Based Lessons from the Teacher’s Perspective

From a teacher’s perspective, the advantages of paper-based lessons are that class time can be used to cover more tasks, concordance lines can be edited to control the focus (i.e., teachers can pre-select salient examples and delete variations or exceptions), the time spent setting up and trouble-shooting computers and software is eliminated, and without a need for IT support, there is no need for a teaching assistant. The main disadvantages are that the exercises are not as interesting, powerful, and motivating; students tend to mechanically underline target structures rather than become active learners; and they are not taught how to use corpus tools.

Conclusion

As described in this paper, findings from the 2008 study show that beginner level students made significant gains in understanding word classes and derivations and in recognizing noun and verb phrases through the use of corpus-based exercises. In each year of the Nihon University Study, students have continued to respond favourably to this approach. Students in the 2009 study reported advantages in both computer-based and paper-based exercises. Although not always as powerful, paper-based corpus exercises can narrow the focus, cover more tasks in the same amount of time and allow students to learn how to use concordances without having to deal with also learning the use of computers and software.

Biodata

Kathryn Oghigian received her Master’s degree in Modern Language Education from the University of British Columbia in 1997. She currently teaches academic reading and technical writing to science and engineering students in the Center for English Language Education for Science and Engineering, Waseda University, Tokyo.

Kiyomi Chujo is an associate professor at the College of Industrial Technology, Nihon University, Japan. Her current research interests are vocabulary learning and the pedagogical

applications of corpus linguistics. She won the JAECS Award in 2008 (Japan Association for English Corpus Studies).

Acknowledgement

This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (B), 2009-2012, 21320107.

References

- Abdon, B., & Raab, R. (2001, September). *Making e-learning work in the Asia Pacific: Lessons learned*. Paper presented at the Information Technology in Regional Areas (ITiRA) Conference, Rockhampton, Queensland, Australia. Retrieved from http://www.sdlearn.net/aprtc/occasional_papers/itirap.htm
- Antconc [Computer software]. Available at <http://www.antlab.sci.waseda.ac.jp/software.html>
- Biber, D. (2009, December). *Using corpus analysis to describe language use*. Seminar paper presentation, Tokyo University of Foreign Studies, Tokyo, Japan.
- 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, 8th Teaching and Language Corpora Conference* (pp.38-44). Lisbon, Portugal: Associação de Estudos e de Investigação Científica do ISLA-Lisbo Fichier.
- Braun, S. (2008). From pedagogically relevant corpora to authentic language learning contents. *ReCALL* 17(1): 47-64.
- Chujo, K. (2003). *Eigo shokyuusha-muke TOEIC-goi 1 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., & Oghigian, K. (2006, July). *Discovering grammar in the beginning-level DDL EFL classroom*. Paper presented at the Seventh International Conference on Teaching and Language Corpora, Paris, France.
- 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, 8th Teaching and Language Corpora Conference* (pp. 65-71). Lisbon, Portugal: Associação de Estudos e de Investigação Científica do ISLA-Lisbo Fichier.
- Collins WordbanksOnline English Corpus Concordance Sampler [Computer program]. Available at <http://www.collinslanguage.com/wordbanks/default.aspx>
- Corpus.BYU.edu [Computer program]. Available at <http://corpus.byu.edu/>
- Huang, L.-S (2008). Using guided, corpus-aided discovery to generate active learning. *English Teaching Forum*, 46(4), 20-27.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge, England: Cambridge University Press.
- Johns, T. (1994). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. In T. Odlin (Ed.), *Perspectives on pedagogical grammar* (pp. 293-313). Cambridge, England: Cambridge University Press.
- Lee, D. (2009). Bookmarks for corpus-based linguists [Website]. Available at <http://tiny.cc/corpora>

- Leech, G. (1997). Teaching and language corpora: A convergence. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 1-23). Edinburgh, Scotland: Addison Wesley Longman Limited.
- Michigan Corpus of Academic Spoken English (MICASE) [Computer program]. Available at <http://micase.elicorpora.info/>
- Nation, P. (2004). A study of the most frequent word families in the British National Corpus. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language* (pp. 3-13). Amsterdam: John Benjamins.
- Norvig, P. (2007, May 31). Theorizing from data: Avoiding the capital mistake. [Video file]. Available from <http://www.youtube.com/watch?v=nU8DcBF-qo4>
- O'Keefe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge, England: Cambridge University Press.
- Paraconc [Computer software]. Available at <http://www.athel.com/mono.html>
- Stevens, V. (1995). Concordancing with language learners: Why? When? What? *CAELL Journal* 6(2), 2-9.
- Tomlinson, B. (Ed.). (1998). *Materials development in language teaching*. Cambridge, England: Cambridge University Press.
- 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). Retrieved from <http://www.asian-efl-journal.com/index.php>
- Utiyama, M., & Isahara, H. (2003). Reliable measures for aligning Japanese-English news articles and sentences. *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics*, 1, 72-79. doi:10.3115/1075096.1075106