Development of a scale to measure learners' perceived preferences and benefits of data-driven learning

Citation

Mizumoto, A., Chujo, K., & Yokota, K. (2016). Development of a scale to measure learners'

    perceived preferences and benefits of data-driven learning. *ReCALL*, 28, 227–246.

    doi:10.1017/S0958344015000208

Development of a scale to measure learners' perceived preferences and benefits of data-driven learning

ATSUSHI MIZUMOTO (mizumoto@kansai-u.ac.jp)
Kansai University, Japan

KIYOMI CHUJO (chuujou.kiyomi@nihon-u.ac.jp)
Nihon University, Japan

KENJI YOKOTA (yokota.kenji@nihon-u.ac.jp)
Nihon University Japan

Abstract

In spite of researchers' and practitioners' increasing attention to data-driven learning (DDL) and increasing numbers of DDL studies, a multi-item scale to measure learners' attitude toward DDL has not been developed thus far. In the present study, we developed and validated a psychometric scale to measure learners' perceived preferences and benefits of DDL for research and pedagogical purposes. First, we created an item pool by referring to open-ended responses from learners; second, the items were pilot tested with target-level learners of English as a foreign language; and third, with item analyses and exploratory factor analysis, the revised version of the questionnaire was prepared. Finally, the questionnaire was administered, and its psychometric properties were examined with confirmatory factor analysis and fit indices. The final phase also included a measure of task values to explore the convergent evidence of the construct validity of the proposed scale. The results suggest that the scale is a valid measure of learners' attitudes toward DDL, with the hypothesized model providing a good fit with the data. We propose that the scale can be used in future studies that utilize the same type of questionnaire research to facilitate further investigation of DDL. Suggestions for further research are also provided.

Keywords: data-driven learning (DDL); learner concordancing; computer use in classroom; attitude; scale development; CALL

1. Introduction

In the course of its development as a discipline, corpus linguistics has had a significant effect on the field of applied linguistics (Hunston, 2002). Its effects can be seen in materials development (e.g., dictionaries, usage manuals, grammar books, and course books), test development (Taylor & Barker, 2008), and the fact that many introductory books on second language acquisition (SLA), language teaching, language testing, and materials development include sections on corpus linguistics or its applied domains (e.g., Loewen & Reinders, 2011; Long & Doughty, 2011; Mackey & Gass, 2012; Shohamy & Hornberger, 2008; Tomlinson, 2013).

Corpus research also has a deep connection with computer-assisted language learning (CALL). Specifically, among the pedagogical applications of corpora such as learning and teaching of vocabulary and phraseology, grammar, pragmatics, writing, reading, speaking, listening, and translation (e.g., Aijmer, 2009; Aston, 2001; Flowerdew, 2012; O'Keeffe, McCarthy & Carter, 2007; Reppen, 2010; Römer, 2010; Sinclair, 2004), direct applications of corpora in which learners themselves get hands-on experience of using a corpus for learning purposes, often with guided tasks or materials, are called "data-driven learning" (henceforth DDL). The term "DDL" was coined by Johns (1990) more than 20 years ago, suggesting that

the "language-learner is also, essentially, a research worker whose learning needs to be driven by access to linguistic data—hence the term 'data-driven learning' (DDL) to describe the approach" (p. 2). It should be noted here we accept a broad definition of DDL (Römer, 2011) involving both hands-on (i.e., direct searches of corpora by learners) and hands-off uses (i.e., direct searches of corpora by teachers, who prepare paper-based materials for learners; so-called "paper-based" DDL).

Research to date suggests that the benefits of DDL include input enhancement from multiple contexts in a concordancer, rich exposure to authentic language use, awareness raising (or noticing) toward patterns and forms, cognitive and meta-cognitive development, improved skills and communicative ability, heightened motivation, student-centeredness, and inductive learning. All of these benefits could result in greater autonomy and life-long learning (Boulton, 2009, 2010; Gilquin & Granger, 2010; Lin & Lee, 2015; O'Sullivan, 2007; C. Yoon, 2011).

In the field of CALL, research papers on DDL have appeared in the major international journals (e.g., Boulton & Pérez-Paredes, 2014; Geluso, 2013). Along with other corpus-integrated activities, the DDL approach has been regarded as a language learning and teaching methodology. It has been mainly used to show concordance lines to students for the purpose of learning and teaching lexico-grammatical or grammar structure with patterns. Especially in the field of English for Academic or Specific Purposes (EAP/ESP), it has been used often as an L2 writing and reference tool (e.g., Friginal, 2013; H. Yoon & Jo, 2014). Although it is generally assumed that DDL works only for advanced learners and needs extensive training (e.g., Kennedy & Miceli, 2001), researchers have found that that is not the case and argued that DDL works equally well for lower-level learners without going through lengthy training in using a concordancer with the aid of teacher-prepared materials (e.g., Boulton, 2009; Chujo & Oghigian, 2012). Boulton (2010), for example, has shown that, for lower-level learners, by utilizing paper-based DDL activities prepared by teachers, DDL in the printed form can have immediate benefits for those learners and can counter potential barriers that inhibit the use of DDL in the classroom.

Recent years have seen a growing body of research that examines the effects of DDL in the classroom. For example, DDL has proved effective in teaching phrasal verbs (Azzaro, 2012), in improving the use of linking adverbials, reporting verbs, and verb tenses in academic writing (Friginal, 2013), and in acquiring lexico-grammatical patterns, shown in improved accuracy and complexity in L2 writing (Huang, 2014). Frankenberg-Garcia (2014) showed that DDL could be useful for both receptive and production purposes. Using paper-based DDL, Smart (2014) put forward positive results that DDL helped learners improve their grammar ability with the passive voice. It should be emphasized that all the studies reported above as examples of those investigating the effectiveness of DDL as a teaching methodology provide evidence that experimental groups with DDL outperformed control groups without DDL, supporting the use of DDL in the classroom over other teaching methods and techniques. Further, there are a number of narrative syntheses on DDL such as Cheng (2010: 320), who concludes that "DDL has been found to be a useful language learning methodology, and there is evidence that learners can indeed benefit from being both language learners and language researchers." However, it is plausible that only advocates of DDL or corpus-based teaching methodology report positive evidence for DDL. Thus, more reliable and quantitative evidence of effectiveness of DDL as a teaching methodology at the meta-analysis level is necessary in order to evaluate the effects of its use objectively.

Meta-analysis is a secondary research methodology to statistically combine research outcomes represented in a quantifiable unit measured with instruments such as tests and psychological scales (e.g., Norris, 2012). With meta-analysis, researchers can put forward more rigorous empirical evidence, as it is an established method to integrate the results from

the primary studies. Recently, Cobb and Boulton (2015) conducted a meta-analysis of DDL studies, including improving writing, learning vocabulary and grammar, reading comprehension, and noticing skills. Their meta-analysis included 21 DDL studies out of 116 (from 1989 to 2012), which reported the essential descriptive statistics such as the number of participants, means, and standard deviations for calculating the effect sizes. Cobb and Boulton reported that, for the between-group contrasts ($k = 13$), the combined effect size was $d = 1.04$, 95% CI [0.83, 1.25], and for the pre-post or within-group contrasts ($k = 8$) was $d = 1.68$, 95% CI [1.36, 2.00]. Both can be regarded as large effect sizes following the criteria of Plonsky and Oswald (2014). From these results, Cobb and Boulton concluded that, especially by comparing the results from meta-analyses of instructed SLA in general (Norris & Ortega, 2000; Spada & Tomita, 2010) and CALL in particular (Grgurović, Chapelle & Shelley, 2013), corpus use in the classroom is effective and results in sizable gains in the outcome measures.

Despite its reported effectiveness, DDL has not become part of mainstream teaching practice. This may be because, as Gilquin and Granger (2010) rightly pointed out, problems and limitations of DDL exist in (a) the logistics, (b) the teacher's point of view, (c) the learner's point of view, and (d) the content of DDL. As for (a) the logistics, we tacitly assume that DDL involves the use of concordancing software in a classroom where computers are available. However, there are a few studies that report the effectiveness of paper-based material as a variant form of DDL activities (e.g., Boulton, 2010; Chujo, Anthony, Oghigian & Uchibori, 2012; Chujo & Oghigian, 2012). Considering that beliefs and self-efficacy play a pivotal role in learning and teaching a second or foreign language (e.g., Dörnyei, 2005), (b) the teacher's point of view and (c) the learner's point of view toward DDL should be of paramount importance, especially when those new to the DDL approach have negative opinions about DDL or have no idea about what it is. It has been reported that teachers do not normally learn how to use corpora in teacher training courses (Granath, 2009; Römer, 2009). At the same time, teachers may not have adequate computer expertise or may find preparing DDL materials time-consuming (Lin & Lee, 2015). Learners share similar difficulties with teachers in utilizing the DDL approach, and without appropriate support, factors such as lack of confidence in using a concordancer, its time-consuming nature, and the difficulty of interpreting concordance lines may impede active use of DDL (Geluso & Yamaguchi, 2014). Yet, studies that attempt to change the viewpoints of teachers or learners toward DDL have not been conducted thus far. Regarding (d) the content of DDL in an EAP or ESP course, although DDL could be used for a communicative learning task (Cresswell, 2007) and such use in a communicative context is recommended (Aston, 2001), mostly it is either geared towards learning and teaching of lexico-grammatical items or is used as an L2 writing reference tool, as noted earlier. The specific content (of the corpus data and syllabuses) and outcome measures in the DDL approach vary from study to study, which may limit the generalizability of the findings reported. Obviously, these concerns need to be further addressed and overcome because research on DDL as a learning and teaching methodology will likely continue, we need to tackle each problem in the DDL approach more thoroughly and systematically.

In this study, of the four problems and limitations listed above, we focused on "the learner's point of view." In many of the empirical studies investigating DDL, attitudes toward the approach are reported and discussed (Boulton, 2009). However, no study to date has measured the learner's point of view (i.e., learner's attitude toward DDL) with a psychometric "scale" (e.g., a Likert scale) with more than two items for each construct (also called multi-item scales, Dörnyei & Taguchi, 2010: 23). The scale score, which can be gained from adding (or averaging) item scores for similar questions, represents an underlying trait such as attitudes, beliefs, and other mental variables. On the other hand, DDL studies have conventionally employed a single-item scale, which supposedly taps into the target construct

with only one item. For instance, H. Yoon and Hirvela (2004) asked students for their opinions (perceptions of the strengths and weaknesses) about using the Collins COBUILD Corpus in L2 writing. Although this study was highly revealing in that it provided an overview of students' opinions about using DDL, only one item was used to measure each construct (e.g., "The corpus is more helpful than a dictionary for my English writing."). As such, the response of a student for a specific item may have been unreliable due to measurement error. As measurement error is virtually always present in any measurement, we need a scale consisting of more than two items rather than one item for more accurate measurement. As an illustrative example, imagine a situation where a researcher intends to measure a learner's grammatical ability with multiple-choice questions. Of course, the researcher would prepare more than one item to measure the learner's grammatical ability, such as accurate use of the target irregular verbs, because measuring it with only one item would certainly entail a measurement error (e.g., resulting from mood, fatigue, and wording of the items). When it comes to questionnaires, researchers often fail to take measurement error into account. As Dörnyei and Taguchi (2010: 23) argue, "the notion of multi-item scales is the central component in scientific questionnaire design, yet this concept is surprisingly little known in the L2 profession." Dörnyei and Taguchi also continue that "because of the fallibility of single items, there is a general consensus among survey specialists that more than one item is needed to address each identified content area" (p. 25). With multi-item scales, it is possible to calculate a reliability coefficient such as Cronbach's α, which reflects that the scale takes measurement error into consideration. Furthermore, it has been reported that "multi-item scales clearly outperform single items in terms of predictive validity" (Diamantopoulos, Sarstedt, Fuchs, Wilczynski & Kaiser, 2012). Note that some researchers are against using Likert scale items as interval scales because they are in fact ordinal (for which adding or averaging item scores is inappropriate). However, it is conventionally accepted that Likert scales could be treated as interval when the wider range of scales are included (Hatch & Lazaraton, 1991).

Although studies employing a questionnaire with a single-item approach could shed light on learners' attitudes toward DDL, for the reasons stated above we need to develop a psychometric scale to further investigate and advance the research on DDL. This is especially true considering the fact that DDL is claimed to foster autonomy and learner-centeredness in learning (Boulton, 2010; O'Sullivan, 2007). With a psychometrically valid and reliable scale for measuring learners' attitudes toward DDL, sophisticated modern statistical analyses can be conducted to further understand the relationship between learners' attitudes toward DDL and other potentially influencing variables in a more scientific manner. Thus, the aim of the current study was to develop and validate a reliable multi-item scale (in the form of a questionnaire) for measuring learners' attitudes toward DDL (i.e., perceived preferences and benefits).

2. Method
2.1. Treatments in previous studies
Before we describe how we developed an item pool for the questionnaire, in this section, we explain the treatments in previous studies, in which we have collected open-ended responses about the perceived preferences and benefits learners feel toward the DDL approach. The second author of this article and her colleagues have been incorporating DDL in their classroom teaching practice since 2004 (Chujo, Anthony, Oghigian & Uchibori, 2012), targeting beginner-level EFL university learners (A1 to A2 in the Common European Framework of Reference for Languages; CEFR, Council of Europe, 2001). The learners' nationality and L1 are Japanese. They are mostly science and engineering majors at a private university in the Tokyo metropolitan area of Japan. DDL applications began as a four-week

(once a week; 90 minutes) activity for learning vocabulary items. They have since been conducted with pedagogical modifications made every year, together with the development of user-friendly web-based bilingual concordancers (Anthony, Chujo & Oghigian, 2011). Since 2006, they have been expanded to 20 weeks (i.e., two semesters of 10 weeks; 90 minutes per week) with DDL syllabuses such as those listed in Table 1.

The DDL syllabuses were developed based on the focus of the course—preparing those beginner-level learners for taking the Test of English for International Communication (TOEIC). In other cases, the focus involved learning rudimentary grammar rules for remedial-level learners (Chujo, Anthony, Oghigian & Yokota, 2013). Also, concordancers and teaching procedures have been refined each year to meet the needs of evolving DDL syllabuses. Thus, we list the 2008 DDL syllabus here as an example and provide a detailed description of the treatment. In this course, each class had a vocabulary component and a DDL grammar component. The vocabulary was grouped into categories such as business, personnel, meetings, and travel, and learners studied a total of 400 words over two semesters. In each class, learners first used a CALL vocabulary program to study twenty vocabulary words. Seven of those twenty words were used as the focal point of the subsequent DDL grammar lesson.

[Table 1 near here.]

Both the grammar and vocabulary items covered in the syllabus used in 2008 (Table 1) were selected following research findings that suggest grammar and vocabulary items to learn to better prepare beginner-level learners for the TOEIC test (for details, see Chujo & Oghigian, 2012).

In teaching grammar and vocabulary items, teachers have utilized bilingual corpora such as a Japanese-English newspaper corpus (Utiyama & Isahara, 2003) for TOEIC preparation courses, and the Corpus of Beginner Level English (Chujo, Anthony, Oghigian & Yokota, 2013) for remedial-level courses aiming at learning junior- or high school-level grammar items. Such bilingual corpora and parallel concordancers enable learners to understand the target language concordance lines and show a context in both languages in the format of Key Word in Context (KWIC), which is a prerequisite for an effective use of DDL (Chujo, Anthony & Oghigian, 2009). The rationale behind using bilingual corpora and parallel concordancers was that L1 translations would help overcome common difficulties (e.g., lack of confidence and difficulty in interpreting concordance lines) that lower-level learners would face in using monolingual concordancers. The following four-stage approach, which was developed in 2008 (Chujo & Oghigian, 2008), has been employed in the classroom DDL activities (Chujo, Anthony, Oghigian & Yokota, 2013, p. 73):

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

In Stage 1, students work in pairs or groups sharing their findings and offering support to each other; consequently, they arrive at hypotheses about the form and usage of a particular lexico-grammatical pattern (Figure 1). In Stage 2, the teacher explains the target items so that

students can confirm if their hypotheses were correct or not. In Stage 3, students are assigned to work on additional practice and consolidation exercises as homework. In Stage 4, students again work together to complete the production practice exercises in class and the teacher gives feedback on the exercises.

The four-stage approach may seem like a traditional class procedure where a teacher-centered, deductive PPP (presentation, practice, and production) method is used. That, however, is not the case. In Stage 1, students play an active role in formulating their hypotheses, exchanging their ideas with peers in pairs or in groups, and discovering grammatical patterns by themselves. This type of interaction helps promote scaffolding in the Zone of Proximal Development (e.g., van Lier, 2004) wherein learners with different levels of proficiency and learning styles can help each other. The teacher's role, therefore, is that of a facilitator at this point, encouraging discovery learning. From Stage 2, explicit explanation takes place, in which the deductive PPP teaching method is used. In this sense, our four-stage approach is a hybrid of an inductive DDL approach and a deductive grammar teaching method, one very similar to the "guided induction" (Flowerdew, 2009; Smart, 2014).

Chujo and Oghigian (2012) report that, in their studies using the DDL approach described above, student gains in the target lexico-grammatical items and overall proficiency, measured with the pre-post test design, have been consistently positive over the years. Moreover, Mizumoto and Chujo (2015) conducted a meta-analysis of their DDL studies with a pre-post design and concluded that the synthesized results, based on the classification of the outcome measures, showed that the DDL approach worked well particularly for learning vocabulary items ($k = 4$, $d = 2.93$, 95% CI [2.19, 3.67]). Other effects were calculated for basic grammar items ($k = 9$, $d = 0.81$, 95% CI [0.69, 0.93]), noun and verb phrases ($k = 14$, $d = 0.86$, 95% CI [0.73, 0.99]), and for proficiency ($k = 5$, $d = 0.40$, 95% CI [0.22, 0.58]). Note that the proposed criteria for d values from pre-post or within-group contrasts are: 0.60 small, 1.00 medium, and 1.40 large (Plonsky & Oswald, 2014).

[Figure 1. An example of a Japanese-English parallel corpus and a sample task for the noun phrases.]

In addition to those sizable gains in the learning outcomes, students continue to report that they enjoyed the DDL approach and found it useful and effective for learning grammar and vocabulary. For example, Figure 2 illustrates the accumulated feedback from students on the perceived effectiveness of DDL activities for grammar and vocabulary learning in the teaching practice of four consecutive years from 2006 to 2010 ($N = 103$).

As Figure 2 shows, a majority of students perceived the DDL activities to learn grammar and vocabulary items as effective, with 64.1% of the students feeling positive about DDL activities in learning grammar and 70.9% in learning vocabulary.

[Figure 2. Accumulated feedback from students on the perceived effectiveness of DDL activities for grammar and vocabulary learning in the classroom from 2006 to 2010 ($N = 103$).]

2.2. Development of the item pool

In their series of classroom applications of DDL stretching over 10 years (described above), the second author of this article and her research team have amassed students' reflective open-ended responses (i.e., perceived preferences and benefits) to questions about DDL at the end of each course. Some of their studies were concerned with organizing these reflective comments into categories and coding them accordingly. Chujo, Anthony and Oghigian (2009) performed a text analysis (i.e., text mining) using these open-ended responses and found that words such as *can, word, usage, variety of, examine, observe, search, immediately* were used more frequently than other words, suggesting that the beginner-level learners in the primary studies appreciate that they can examine and observe a variety of examples of authentic usage immediately and easily. In the process of writing questionnaire items, we referred to the original contexts in which those words were used. Although negative keywords such as *eyestrain, lack of time*, and *initial settings* and open-ended responses such as "*It first required time to get used to the search operation*" were also reported (Chujo, Anthony & Oghigian, 2009), we decided to focus on the perceived preferences and benefits of DDL, and did not include the drawbacks. This is because, in unidimensionality scale development (i.e., measuring only one overarching construct in an instrument), it would be theoretically and statistically unsupportive to attempt to measure mixed constructs. In other words, we could not create a one-size-fits-all measurement instrument in theory. While recognizing the potential and pedagogical value of including drawbacks in the questionnaire items, we chose statistical and methodological rigor over practical utility, which is inevitable in a scale development study like this one.

By referring to these open-ended responses and items or categories in previous studies (e.g., H. Yoon & Hirvela, 2004), we developed a list of questionnaire items. The wording of the items was examined and modified where necessary by the authors of this article. In total, 18 items were developed as a result of this procedure and included in the questionnaire (see Table 3 for items).

2.3. Field-testing the questionnaire

The questionnaire was field tested with 267 university EFL learners (science and engineering majors, 226 males and 41 females, aged 18–20) at a private university in Japan. Their proficiency measured with the TOEIC Bridge test was beginner level ($n = 255$, $M = 131.73$, $SD = 14.08$; the score is roughly equivalent to 350 in the TOEIC test), which, according to Educational Testing Service (2007), is classified as "Basic User" (A1 to A2) in the CEFR. We focused on the learners at this proficiency level because approximately 80% of the target Japanese EFL learners (university undergraduates) are at that level (Tono & Negishi, 2012). Although the proficiency measures of the respondents in the previous studies were not available, the authors confirmed, based on observations and in-house examinations (and also from the fact that the DDL intervention had been conducted at the same university for the past ten years), that the participants who responded to the questionnaire had about the same proficiency level as the samples in the previous studies.

The questionnaire was administered after a three-month DDL intervention (according to the DDL syllabus and procedures explained in Chujo, Anthony, Oghigian & Uchibori, 2012) as part of a compulsory English course at their university. The participants responded to the 18-item questionnaire on a 6-point scale: 1 (not at all true of me), 2 (not true of me), 3 (somewhat not true of me), 4 (a little true of me), 5 (mostly true of me), and 6 (very true of me), according to the degree of perceived preferences and benefits of DDL. The decision of using a 6-point scale (an even number without the "middle" option) was based on a suggestion by Dörnyei and Taguchi (2010: 28).

After pilot-testing the questionnaire, item analyses were carried out consisting of the following steps: (a) examining the item-total correlations to determine whether the figures were over 0.3, which is a suggested criterion for this index of item analysis (Wintergerst, DeCapua & Itzen, 2001: 391); (b) using exploratory factor analysis to investigate which items belong together (i.e., construct validity); (c) scrutinizing Cronbach's α levels to verify the internal consistency of the subscales.

R version 3.1.0 (R Core Team, 2014) was used for all the statistical analyses in this study. In exploratory factor analysis, maximum likelihood extraction with promax rotation was performed. To decide on the number of factors, we first looked at the scree plot to check any distinctive slope between factors, and retained those with the eigenvalues greater than 1.0, which is an accepted criterion for extracting factors. Then, items showing factor loadings above 0.4 on only one factor were adopted. After going through these screening procedures, 16 of the 18 items were retained for the final version of the questionnaire.

Table 2 shows the result of the exploratory factor analysis. Two factors were extracted and, through an extensive discussion among the authors, they were named after advantages of DDL reported in the literature (e.g., Johansson, 2009; Kennedy & Miceli, 2001; O'Sullivan, 2007). Table 3 lists the questionnaire items.

The first factor was "Clarity," which consists of items such as "I can see the target sentences in real use" (Item 05) and "I can see many sentences that include the target structure" (Item 01). These items intend to measure the advantages of DDL in clarifying the authentic use of target structures. The second factor was "Autonomy" with items such as "this type of learning is not passive but active" (Item 12) and "I can search for and learn target sentences independently" (Item 11). These items seem to represent the extent to which the learners embrace the autonomy of learning in the DDL approach. Internal consistency reliability (Cronbach's α) of the items representing each factor was moderately high with .91 for Clarity and .81 for Autonomy. The Pearson correlation coefficient between Clarity and Autonomy was .60, which is reasonably high, and thus it can be suggested that these two subscales (i.e., Clarity and Autonomy) successfully measure similar constructs under an overarching theme, "perceived preferences and benefits of DDL."

[Table 2 near here.]

[Table 3 near here.]

2.4. *Administering the questionnaire*

In order to validate the factor structure, the 16-item questionnaire targeting the perceived preferences and benefits of DDL was administered to another group of 147 EFL learners. The major and proficiency level of the participants were about the same as those of the learners in the pilot study (science and engineering majors, 123 males and 24 females, aged 18–20; the TOEIC Bridge test score, $n = 140$, $M = 131.00$, $SD = 15.82$). We administered the questionnaire to the learners who take the same courses that incorporate a DDL approach at the same university in the year following the pilot study. Like the pilot study, the questionnaire was administered after a three-month DDL intervention as part of a compulsory English course at their university. The participants again responded to the 16-item questionnaire on a 6-point scale, from 1 (not at all true of me) to 6 (very true of me), according to the degree of perceived preferences and benefits of DDL.

The construct validity of the questionnaire was examined with confirmatory factor analysis (CFA) because the decisions about the factor models were made a priori (from the pilot study). In order to ensure the unidimensionality (i.e., measuring only one underlying trait) of the subscales of the questionnaire, CFA was conducted for each subscale, Clarity and Autonomy. Based on a criterion for higher convergent validity (Hair, Black, Babin, Anderson & Tatham, 2006: 777), items with standardized loading estimates higher than 0.5 were used for further analysis. This procedure left five items for both Clarity and Autonomy measures. In structural equation modeling (SEM), an umbrella term for a collection of methods including CFA, a set of observed variables (i.e., the questionnaire items in this study), latent variables (i.e., underlying traits that can be represented with the subscales of the questionnaire – Autonomy and Clarity, in this study), and measurement errors can be modeled simultaneously. Because SEM tests whether a hypothesized model is consistent with the observed data, several fit indices are examined to evaluate the overall fit of the structural model. In this study, we checked the comparative fit index (CFI), Tucker-Lewis index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). The acceptable criterion value for each fit index is: CFI and TLI > .90, RMSEA and SRMR < 0.08. Although the chi-square value, with which a familiar p value can be computed, is the basis of these indices, the evaluation of goodness of fit with chi-square value is not recommended because it is sensitive to sample size and other factors (Hair et al., 2006). Thus, it is a common practice in SEM to examine and compare several goodness-of-fit indices, which do not come with p values. The one-factor model provided a good fit to the data for both the subscales of Autonomy (CFI = .97, TLI = .94, RMSEA = 0.09 [90% CI = 0.01–0.17], SRMR = .04) and Clarity (CFI = 1.00, TLI = 1.00, RMSEA = 0.00 [90% CI = 0.00–0.11], SRMR = .02).

In addition to the questionnaire asking about the perceived preferences and benefits of DDL, six items purported to measure task values, which can be defined as individual beliefs about the relative worth of tasks, were prepared (Table 4) by referring to items such as those in the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich, Smith, Garcia & McKeachie, 1993) and administered to the same participants. The rationale behind measuring the task values is that we can expect a relationship between the perceived preferences and benefits of DDL and the task values. That is, if one favors and appreciates the benefits of DDL, it is likely that one's task values will increase, and probably vice versa as well (i.e., high task values would correspond to higher degrees of perceived preferences and benefits of DDL). This would provide us with convergent evidence of the construct validity of the proposed scale (i.e., external aspect of validity in Messick, 1995). With CFA, the one factor model of task values with six items (see Appendix B) showed a good fit to the data (CFI = .97, TLI = .95, RMSEA = 0.12 [95% CI = 0.07–0.17], SRMR = .04). The relationship among the perceived preferences and benefits of DDL and task values was also explored with a path analysis using SEM and its model fits. For the purpose of transparent sharing of data and their analyses, the data and R codes used in this study are available online (http://mizumot.com/files/ReCALL_DDL.html).

[Table 4 near here.]

3. Results and Discussion
Internal consistency reliability (Cronbach's α) and descriptive statistics of all the scales are displayed in Table 5. Cronbach's α reliability coefficients for the three scales, Clarity, Autonomy, and Task Value, were relatively high. These three scales showed reasonably high

correlation coefficients (i.e., all values above .50). According to Dörnyei (2007: 223), "in applied linguistics research we can find meaningful correlations of as low as 0.3–0.5 . . . and if two tests correlate with each other in the order of 0.6, we can say that they measure more or less the same thing." Furthermore, a more recent meta-analysis of L2 studies by Plonsky and Oswald (2014) suggests that a correlation coefficient of .60 can be considered a large effect (i.e., a strong relationship). Applying these criteria, we can argue that the two subscales of perceived preferences and benefits of DDL (i.e., Clarity and Autonomy) had a strong correlation ($r = .58$). Each of these two subscales showed high correlations with Task Value (Clarity and Task Value, $r = .58$; Autonomy and Task Value, $r = .62$). That is, the more positive the learners' attitudes toward perceived preferences and benefits of DDL (i.e., Clarity and Autonomy) are, the higher their Task Value would be, and vice versa. These patterns exactly reflected the hypothesized relationship between the perceived preferences/benefits of DDL and task values.

[Table 5 near here.]

Figure 3 shows the result of SEM (structural equation modeling), which explored the relationship between the perceived benefits of DDL and task values. The model assumes that the higher-order factor (i.e., Perceived Benefits of DDL) affects the two factors of DDL (i.e., Clarity and Autonomy), and that the higher-order factor correlates with Task Value. In Figure 3, the variables are illustrated with boxes and circles. Boxes show that they are observed variables (i.e., questionnaire items) and circles are latent variables (i.e., constructs or underlying traits that affect the response of each questionnaire item). The arrows show the standardized path coefficients with a possible range of -1 to 1. Those paths represent the strength of the relationship between variables, with the higher values indicating stronger relationships. The directions of arrows indicate the hypothesized causality within the model. For each item and two factors (Clarity and Autonomy), measurement errors, which cannot be explained with the hypothesized model, are also displayed in Figure 3. Because SEM can include measurement error in the model in this way, it is a known advantage of using SEM in data analysis.

The fit indices revealed that the hypothesized model provided an adequate fit to the data (CFI = .93, TLI = .92, RMSEA = 0.08 [95% CI = 0.07–0.10], SRMR = .06). Because examining the higher-order factor (latent variable) and its relationship with other variables (in this case, Task Value) and evaluating the model fit indices are not possible with normal correlation analyses (Table 5), the hypothesized model in Figure 3 provides stronger statistical evidence that Perceived Benefits of DDL and Task Value positively affect each other (with the standardized correlation coefficient of .82).

[Figure 3. SEM result of the relationship between the perceived benefits of DDL and Task Value]

Taken together, with adequate internal consistency reliability coefficients, high inter-correlations, and established convergent evidence of construct validity (i.e., the expected relationship between the subscales and task values), the two subscales in the questionnaire serve as a valid and reliable measure of the learners' perceived preferences and benefits of DDL. This study is unique in that it provides a research and pedagogical instrument for

assessing the effectiveness of DDL (and possibly expanded into CALL in general) from the viewpoints of learners. As such, the questionnaire developed in this study will be useful for systematic inquiry into learners' attitudes toward DDL.

Although we can justifiably claim that we have achieved our initial research purpose of developing a psychometrically sound instrument for assessing learners' perceived preferences and benefits of DDL, the current study has a few limitations that should be acknowledged. First, the participant groups here (in both pilot and main studies) were homogeneous, composed of false-beginner level learners. Second, the study took place only in one university in Japan. It can be assumed that the socio-cultural context the learners are in will influence the way they perceive DDL. Therefore, studies on the effectiveness of DDL with more proficient learners, at different institutions, and in second or other foreign language learning settings, should be conducted to generalize the applicability of the instrument. Third, the questionnaire items were created based on a "guided induction approach" (Flowerdew, 2009; Smart, 2014), where learners use self-explanatory guided worksheets created by teachers and mainly engage in exploring lexico-grammatical items in a classroom. As pointed out by Johansson (2009: 41) and Flowerdew (2012: 197), though DDL is usually associated with an inductive approach, in actual DDL classroom activities, a "deductive" DDL approach is often used. Such use of DDL is in stark contrast to inductive DDL, where advanced learners consult the concordance lines themselves with much less help from the teacher (Boulton, 2010; Cresswell, 2007; Flowerdew, 2009). Because the scale developed in the current study is mainly for the former purposes (i.e., deductive DDL) and it is possible that given different tasks, learners would respond to each item differently, caution should be exercised when researchers employ the scale for the latter purposes (i.e., inductive DDL). Fourth, the developed scale deals with learners' perceived preferences and benefits of DDL, and problems or difficulties in corpus use (H. Yoon & Hirvela, 2004) were not included in the constructs it measures. This is, as described earlier in this paper, because it is not possible to create a one-size-fits-all measurement instrument, and we chose statistical and methodological rigor accordingly. Thus, those researchers who are interested in negative aspects of DDL will need to devise a similar scale to investigate the topic precisely and in more detail with a psychometrically based multi-item scale.

These limitations being pointed out, findings from the current study have several implications for research and practice involving DDL and CALL in general. First, the questionnaire could serve as a tool to investigate the relationship between the attitude toward DDL and outcomes (i.e., test scores) or other learner variables such as learning styles, learning strategies, motivation, and self-regulation (e.g., Dörnyei, 2005) with much more precision in measurement by using this new instrument. In this sense, although the current study fits previous research on learners' assessment of the benefits of DDL, it does not confirm or contradict previous findings in a directly comparable manner. In other words, it opens up a new, sophisticated, and more desirable assessment dimension for researchers and practitioners involved in DDL. Second, the instrument developed in this study can be used for the evaluation of DDL practice, which in the past could be achieved with test scores or open-ended responses from learners. With this instrument, it is possible to quantify attitudes toward DDL use in language learning and teaching. Third, related to the second implication, meta-analysis or research synthesis of DDL in terms of how learners perceive it is now feasible because the questionnaire developed is a quantifiable, valid, and reliable scale of the trait (i.e., learners' perceived preferences and benefits of DDL). With the questionnaire developed in this study, it will also be possible to synthesize and evaluate learners' attitude toward DDL in a quantifiable fashion. As can be illustrated in these implications, the newly developed scale could have the potential to greatly expand the research possibilities of DDL and CALL in general.

4. Conclusion

The primary purpose of this study was to develop and validate a psychometrically sound scale to measure learners' perceived preferences and benefits of DDL in order to provide a valid and reliable instrument for research and pedagogical purposes. Through the development phases, it was confirmed that the instrument possesses excellent psychometric properties as a measure of learners' perceived preferences and benefits of DDL.

Based on the findings, we suggest that researchers who investigate the role of DDL as a teaching methodology can now utilize the psychometrically justifiable multi-item scale developed in the current study. Researchers can also study the relationship between the subscales of the developed questionnaire and other variables (e.g., autonomy), which can allegedly be fostered by engaging in DDL activities.

Given that the research momentum of DDL is likely to continue, we believe the scale developed in this study holds promise as a valid and reliable measure of learners' attitude toward DDL, and more studies using this scale type will be conducted.

References

Aijmer, K. (ed.) (2009) *Corpora and language teaching*. Amsterdam/Philadelphia: John Benjamins.

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

Aston, G. (ed.) (2001) *Learning with corpora*. Houston, TX: Athelstan.

Azzaro, G. (2012) Phrasal verbs through DDL. *Journal of Theories and Research in Education*, **7**(2): 1–24.

Boulton, A. (2009) Testing the limits of data-driven learning: Language proficiency and training. *ReCALL*, **21**(1): 37–54. doi:10.1017/S0958344009000068

Boulton, A. (2010) Data-driven learning: Taking the computer out of the equation. *Language Learning*, **60**(3): 534–572. doi:10.1111/j.1467-9922.2010.00566.x

Boulton, A. and Pérez-Paredes, P. (2014) ReCALL special issue: Researching uses of corpora for language teaching and learning [Editorial]. *ReCALL*, **26**(2): 121–127. doi:10.1017/S0958344014000068

Cheng, W. (2010) What can a corpus tell us about language teaching? In O'Keeffe, A. and McCarthy, M. (eds.), *The Routledge handbook of corpus linguistics*. London, England: Routledge, 319–332.

Chujo, K., Anthony, L. and 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 Mahlberg, M., González-Díaz, V. and Smith, C. (eds.), *Proceedings of 5th Corpus Linguistics Conference 2009, University of Liverpool, UK*. http://ucrel.lancs.ac.uk/publications/cl2009

Chujo, K., Anthony, L., Oghigian, K. and Uchibori, A. (2012) Paper-based, computer-based, and combined data-driven learning using a web-based concordancer. *Language Education in Asia*, **3**(2): 132–145. http://www.camtesol.org/Download/LEiA_Vol3_Iss2_2012/LEiA_V3_I2_02_Chujo_et_al_Paper_Computer_and_Combined_DDL.pdf

Chujo, K., Anthony, L., Oghigian, K. and Yokota, K. (2013) Teaching remedial grammar through data-driven learning using AntPConc. *Taiwan International ESP Journal*, **5**: 65–90. http://tespaj.tespa.org.tw/index.php/TESPJ/article/view/48

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

Chujo, K. and Oghigian, K. (2012) DDL for EFL beginners : A report on student gains and views on paper-based concordancing and the role of L1. In Thomas, J. and Boulton, A. (eds.), *Input, Process and Product: Developments in Teaching and Language Corpora*. Brno, Czech Republic: Masaryk University Press, 169–182.

Cobb, T. and Boulton, A. (2015) Classroom applications of corpus analysis. In Biber, D. and Reppen, R. (eds.), *The Cambridge handbook of corpus linguistics*. Cambridge University Press, 478–497.

Council of Europe. (2001) *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge University Press.

Cresswell, A. (2007) Getting to "know" connectors? Evaluating data-driven learning in a writing skills course. *Language and Computers*, **61**(1): 267–287. http://core.kmi.open.ac.uk/download/pdf/11070142.pdf

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P. and Kaiser, S. (2012) Guidelines for choosing between multi-item and single-item scales for construct measurement: A predictive validity perspective. *Journal of the Academy of Marketing Science*, **40**(3): 434–449. doi:10.1007/s11747-011-0300-3

Dörnyei, Z. (2005) *The psychology of the language learner: Individual differences in second language acquisition. Language*. Mahwah, NJ: Lawrence Erlbaum Associates.

Dörnyei, Z. (2007) *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.

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

Educational Testing Service. (2007) *Mapping the TOEIC and TOEIC Bridge tests on the Common European Framework of Reference for Languages*. http://www.ets.org/s/toeic/pdf/toeic_cef_mapping_flyer.pdf

Flowerdew, L. (2009) Applying corpus linguistics to pedagogy: A critical evaluation. *International Journal of Corpus Linguistics*, **14**(3): 393–417. doi:10.1075/ijcl.14.3.05flo

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

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

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

Geluso, J. (2013) Phraseology and frequency of occurrence on the web: native speakers' perceptions of Google-informed second language writing. *Computer Assisted Language Learning*, **26**(2): 144–157. doi:10.1080/09588221.2011.639786

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

Gilquin, G. and Granger, S. (2010) How can DDL be used in language teaching? In O'Keeffe, A. and McCarthy, M. (eds.), *The Routledge handbook of corpus linguistics*. London, England: Routledge, 359–369.

Granath, S. (2009) Who benefits from learning how to use corpora? In Aijmer, K. (ed.), *Corpora and language teaching*. Amsterdam/Philadelphia: John Benjamins, 47–65.

Grgurović, M., Chapelle, C. A. and Shelley, M. C. (2013) A meta-analysis of effectiveness studies on computer technology-supported language learning. *ReCALL*, **25**(2): 165–198. doi:10.1017/S0958344013000013

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E. and Tatham, R. L. (2006) *Multivariate data analysis* (6th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Hatch, E. and Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. Boston, MA: Heinle & Heinle.

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

Hunston, S. (2002) *Corpora in applied linguistics*. Cambridge University Press.

Johansson, S. (2009) Some thoughts on corpora and second-language acquisition. In Aijmer, K. (ed.), *Corpora and language teaching*. Amsterdam/Philadelphia: John Benjamins, 33–44.

Johns, T. (1990) From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *CALL Austria*, 10, 14–34.

Johns, T. (1991) From print out to handout: Grammar and vocabulary teaching in the context of data-driven learning. In Johns, T. and King, P. (eds.), *Classroom Concordancing. ELR Journal 4:* 27–46.

Kennedy, C. and Miceli, T. (2001) An evaluation of intermediate students' approaches to corpus investigation. *Language Learning & Technology*, **5**(3), 77–90. http://www.llt.msu.edu/vol5num3/pdf/kennedy.pdf

Loewen, S. and Reinders, H. (2011) *Key concepts in second language acquisition*. Basingstoke, England: Palgrave Macmillan.

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

Long, M. H. and Doughty, C. J. (eds.) (2011) *The handbook of language teaching*. Oxford, England: John Wiley & Sons.

Mackey, A. and Gass, S. M. (eds.) (2012) *Research methods in second language acquisition: A practical guide*. Chichester, England: John Wiley & Sons, Ltd.

Messick, S. (1995) Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, **50**(9): 741–749. doi:10.1037/0003-066X.50.9.741

Mizumoto, A. and Chujo, K. (2015). A meta-analysis of data-driven learning approach in the Japanese EFL classroom. *English Corpus Studies,* **22**: 1–18. http://mizumot.com/files/ecs2015.pdf

Norris, J. M. (2012) Meta-analysis. In Chapelle, C. A. (ed.), *The Encyclopedia of Applied Linguistics.* Malden, MA: John Wiley & Sons, 3653–3662. doi:10.1002/9781405198431.wbeal0761

Norris, J. M. and Ortega, L. (2000) Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, **50**(3): 417–528. doi:10.1111/0023-8333.00136

O'Keeffe, A., McCarthy, M. and Carter, R. (2007) *From corpus to classroom: Language use and language teaching*. Cambridge University Press.

O'Sullivan, Í. (2007) Enhancing a process-oriented approach to literacy and language learning: The role of corpus consultation literacy. *ReCALL*, **19**(3): 269–286. doi:10.1017/S095834400700033X

Pintrich, P. R., Smith, D. A. F., Garcia, T. and McKeachie, W. J. (1993) Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ).

Educational and Psychological Measurement, **53**(3): 801–813. doi:10.1177/0013164493053003024

Plonsky, L. and Oswald, F. L. (2014) How big is "big"? Interpreting effect sizes in L2 research. *Language Learning*, **64**(4): 878–912. doi:10.1111/lang.12079

R Core Team. (2014) R: A language and environment for statistical computing (Version 3.1.1) [Computer software]. Vienna, Austria. http://www.r-project.org/

Reppen, R. (2010) *Using corpora in the language classroom*. Cambridge: Cambridge University Press.

Römer, U. (2009) Corpus research and practice: What help do teachers need and what can we offer? In Aijmer, K. (ed.), *Corpora and language teaching*. Amsterdam/Philadelphia: John Benjamins, 83–98.

Römer, U. (2010) Using general and specialized corpora in English language teaching: Past, present and future. In Campoy-Cubillo, M. C., Bellés Fortuño, B. and Gea-Valor, M. L. (eds.), *Corpus-based approaches to English language teaching*. London, England: Continuum, 18–38.

Römer, U. (2011) Corpus research applications in second language teaching. *Annual Review of Applied Linguistics*, **31**: 205–225. doi:10.1017/S0267190511000055

Shohamy, E. and Hornberger, N. H. (eds.) (2008) *Encyclopedia of language and education, Volume 7: Language testing and assessment*. (2nd ed.). New York, NY: Springer.

Sinclair, J. M. (ed.) (2004) *How to use corpora in language teaching*. Amsterdam/Philadelphia: John Benjamins.

Smart, J. (2014) The role of guided induction in paper-based data-driven learning. *ReCALL*, **26**(2): 184–201. doi:10.1017/S0958344014000081

Spada, N. and Tomita, Y. (2010) Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, **60**(2): 263–308. doi:10.1111/j.1467-9922.2010.00562.x

Taylor, L. and Barker, F. (2008) Using corpora for language assessment. In Shohamy, E. and Hornberger, N. H. (eds.), *Encyclopedia of language and education, Volume 7: Language testing and assessment* (2nd ed.). New York, NY: Springer, 241–256.

Tomlinson, B. (ed.) (2013) *Developing materials for language teaching* (2nd ed.). London, England: Bloomsbury.

Tono, Y. and Negishi, M. (2012) The CEFR-J: Adapting the CEFR for English language teaching in Japan. *Framework & Language Portfolio (FLP) SIG Newsletter*, **8**: 5–12. http://www.tufs.ac.jp/ts/personal/tonolab/cefr-j/Tono%26Negishi2012forJALT_FLPSIG.pdf

Utiyama, M. and Isahara, H. (2003) Reliable measures for aligning Japanese-English news articles and sentences. In Hinrichs, E. and Roth, D. (eds.), *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics*. Sapporo, Japan: The Association for Computational Linguistics, 72–79. http://acl.ldc.upenn.edu/acl2003/main/index.html

van Lier, L. (2004) *The ecology and semiotics of language learning*. Boston, MA: Kluwer Academic.

Wintergerst, A. C., DeCapua, A. and Itzen, R. C. (2001) The construct validity of one learning styles instrument. *System*, **29**(3): 385–403. doi:10.1016/S0346-251X(01)00027-6

Yoon, C. (2011) Concordancing in L2 writing class: An overview of research and issues. *Journal of English for Academic Purposes*, **10**(3): 130–139. doi:10.1016/j.jeap.2011.03.003

Yoon, H. and Hirvela, A. (2004) ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing*, **13**(4): 257–283. doi:10.1016/j.jslw.2004.06.002

Yoon, H. and Jo, J. W. (2014) Direct and indirect access to corpora: An exploratory case study comparing students' error correction and learning strategy use in L2 writing. *Language Learning & Technology*, **18**(1): 96–117. http://llt.msu.edu/issues/february2014/yoonjo.pdf

Table 1
A DDL Syllabus Used in 2008

		Spring Semester		Fall Semester
		Grammar & Vocabulary		Grammar & Vocabulary
1	Word classes	Business 1: access, assembly, commercial, competitive, decline, distribute, efficiency, exhibit, export, franchise, import, inventory, invoice, negotiation, organization, product, quality, regarding, safety, sponsor	VP: intransitive & transitive verbs	Business: company, corporate, corporation, cost, development, environment, expand, expansion, finance, firm, incorporate (INC), labour, law, major, manufacturer, profit, protect, protection, regional, trade
2	Derivations and inflections	Business 2: audit, contract, decrease, economic, equipment, facility, headquarters, industry, location, machinery, manufacture, operation, opportunity, package, provide, region, research, site, strategy, unit	VP: intransitive & transitive verbs SVOO	Personnel 1: accept, advertisement (ad, ads), agreement, benefit, candidate, degree, department, detail, division, employee, employment, excellent, experience, promotion, range, recommend, remain, require, submit
3	Count & non-count nouns	Personnel 1: applicant, apply, associate, boss, career, chief, colleague, employer, executive, management, manager, offer, personnel, promote, requirement, resume, retire, strike, technical, vice	VP: gerunds, to-infinitives, that-clauses	Personnel 2: clerk, director, editor, engineer, force, individual, manage, official, operate, operator, president, publisher, reception, relation, representative, responsibility, responsible, service, staff, various
4	NP: Art + Adj + N	Personnel 2: accountant, architect, assistant, auditor, award, client, consultant, former, full-time, lawyer, overtime, professor, receptionist, secretarial, secretary, security, skilled, supervise	VP: gerunds, to-infinitives, that-clauses, wh-clauses	Meetings: activity, appropriate, attend, bid, claim, committee, complaint, consider, deal, decision, encourage, improvement, indicate, intend, meeting, organize, original, prepare, project
5	NP: Adj + N	Meetings: approve, arrange, conference, confirm, convention, copy, demonstrate, discuss, document, measure, postpone, procedure, proposal, propose, regulation, request, revise, seminar,	VP: gerunds, to-infinitives, that-clauses, wh-clauses	Marketing: affect, analysis, analyze, approximately, availability, case, cause, demand, economy, estimate, factor, field, figure, forecast, inflation, marketing, predict, prefer, quarterly, review
6	NP: Art + Adj + N	Travel: accommodate, agency, agent, arrangement, arrival, baggage, check-out, delay, departure, extend, guide, luggage, pack, rental, reservation, single, terminal, transit, valid, visa	VP: that-clauses	Office Work: according, announcement, application, area, code, complete, complex, emergency, enclose, express, extension, fax, file, form, forward, installation, maintenance, suit, supply, update
7	NP: Art + Adj + N	Money: account, adjustment, capital, coverage, credit, currency, earnings, fund, insurance, invest, investor, loan, payroll, PIN, policy, reimbursement, salary, savings, statement, tax	VP: wh-clauses	Money: additional, amount, bank, bill, budget, cash, charge, check, deposit, earn, expense, fee, financial, investment, million, outstanding, payment, rate, secure, stock
8	NP : followed by PP	Buying & Selling: assure, brochure, buyer, cancel, chain, customer, defective, discount, display, mail-order, merchandise, minimum, order, purchaser, refund, replacement, retailer, subscription	VP: intransitive verbs + complement	Purchasing: brand, clothing, consumer, delivery, examine, expensive, feature, function, furniture, guarantee, handle, item, obtain, price, purchase, receipt, receive, regular, sale, shipment
9	NP: followed by to-infinitives	Time: advance, ahead, annual, current, due, frequently, immediately, lately, monthly, present, previous, prior, promptly, quarter, rapidly, recently, temporary, upcoming, urgent, weekly	VP: to-infinitives (expect him to do)	Daily Life: alarm, apartment, appointment, contain, convenience, daily, electrical, estate, household, list, local, meal, medical, membership, nearby, personal, power, property, remove, theatre
10	NP: followed by -ed, -ing	Daily Life: appliance, beverage, cater, detergent, downtown, electricity, envelope, furnish, gas, identification, inspection, install, newsstand, register, registration, rent, replace, resident, suspend	VP: adverb location	Transportation: airport, automobile, available, concern, construction, crowd, expect, fare, information, license, occur, officer, passenger, permit, repair, reserve, route, seat, traffic, transportation

Note. NP stands for noun phrases; VP stands for verb phrases. Reprinted with permission of Chujo, Anthony and Oghigian (2009).

Table 2

Result of Exploratory Factor Analysis (n = 267)

Item	Factor name	Factor 1	Factor 2	α	
Item 01		.92	-.20		
Item 02		.78	-.03		
Item 03		.77	-.11		
Item 04		.74	.00		
Item 15	Factor 1:	.70	.12		
Item 06	Clarity	.64	.12	.91	
Item 17		.62	.12		
Item 08		.57	.09		
Item 09		.51	.26		
Item 10		.50	.23		
Item 11		-.01	.77		
Item 12		.00	.77		
Item 13	Factor 2:	-.24	.72		
Item 14	Autonomy	-.05	.69	.81	
Item 15		.04	.51		
Item 16		.21	.50		
Correlation matrix		Factor 1	Factor 2	*M*	*SD*
Factor 1: Clarity		—		4.61	0.90
Factor 2: Autonomy		.60	—	4.02	0.98

Note. Items are renumbered for clarity.

Table 3

Questionnaire Items for Learners' Perceived Preferences and Benefits of DDL (Originally in Japanese)

The advantage of DDL is:

Factor 1: Clarity	
Item 01	I can see the target sentences in real use.
Item 02	I can see many sentences that include the target structure.
Item 03	It shows many frequently used example sentences.
Item 04	I can visualize the practical usage.
Item 05	It shows the context where the words are often used.
Item 06	I can see a large number of English sentences easily.
Item 07	I can see many more example sentences than in a dictionary.
Item 08	I can get to see Japanese translations.
Item 09	I was able to understand in what meaning the word is used.
Item 10	I can discover a usage I did not know.
Factor 2: Autonomy	
Item 11	This type of learning is not passive but active.
Item 12	I can search for and learn target sentences independently.
Item 13	It is different from traditional or regular English learning.
Item 14	I can use software I have never used before.
Item 15	We don't use English textbooks; instead, we use computers.
Item 16	I start to think about what part of speech words belong to.
Excluded in the exploratory factor analysis	
Item 17	Words are displayed in an organized manner.
Item 18	I can visualize the various word forms such as inflections and derivations.

Table 4

Questionnaire Items for Measuring Learners' Task Values (Originally in Japanese)

Through the tasks in the classroom, I feel:

Item 01	I was able to improve my English ability.
Item 02	They were useful for grammar and vocabulary learning.
Item 03	The learned grammar and vocabulary were easily fixed in memory.
Item 04	The activities were enjoyable.
Item 05	I was able to understand the grammar and vocabulary items I did not know.
Item 06	They were helpful in understanding the target grammar and vocabulary items.

Table 5

Correlation Coefficients and Descriptive Statistics ($n = 147$)

Scale	Clarity	Autonomy	Task Value	No. of Items	M	SD	α
Clarity	—	.68	.63	5	4.64	0.93	.87
Autonomy	.58	—	.72	5	4.14	0.96	.80
Task Value	.58	.62	—	6	3.90	1.03	.92

Note. Correlation between scale scores (below diagonal); SEM result (above diagonal). Possible range of item response (M): 1 to 6.

Task: Search "**a * opportunity for**". Find what word comes before **opportunity** and write them below.

```
the current problem presented a good opportunity for finally resolving the reclama
ency chaos could be viewed as a good opportunity for the globalization of the yen
mp into plural parties offers a good opportunity for the party to purge itself of
na said. The meetings provide a good opportunity for the government to listen to
ng a vote, the contest offers a good opportunity for us to think about who will b
onal Press Club on Monday was a good opportunity for us to learn whether each par
ince praised the festival as a great opportunity for
is sense Rao's visit will be a great opportunity for Japanese and Indian leaders
Haiti at least has been given a new opportunity for self-government and democracy
Algerian elections afforded a unique opportunity for radical Islam to enter the
```

```
今回 が 、 諫早 問題 で 巻き返す 最大 の 好機 」 と 意気込む 。 輸入 米 は 政府 が 一元
に 呼び込み 、 円 の 国際 化 に つなげる 好機 」 という 側面 も あり 、 邦銀 や 日本 の
を 振る 。 保守 陣営 が 多党 化 し た この 機会 を 、 同 党 は 、 金権 体質 など の 体質
労働 界 も 政策 を 世論 に アピール する 絶好 の 舞台 と なって いる 。 自衛隊 が 治安
が 日本 の 顔 に ふさわしい か を 考える 機会 に なる 。 また 、 三月 一 日 から は 、
会 は 、 その 一端 を 明らか に する 良い 機会 と なった 。 むしろ 大量 の 赤字 国債 発
ない 規模 で 日本 文化 を 紹介 する この 機会 を 大切 に したい 」 と スピーチ した
いて 、 率直 な 意見 の 交換 を する よい 機会 で ある 。 30 歳 代 から 50 歳 代 を
少なくとも 自治 と 民主 主義 へ の 新た な 機会 を 与え られ た 。 —— 日本 の 資本 市場
的 に 政治 の 本流 に 加わる ユニーク な 好機 を 与える もの だった 。 残された 時間
```

	Determiner	Adjective	Head Noun		Post-modifier
1	( )	( )	opportunity	for	finally resolving …
2	( )	( )	opportunity	for	the globalization …
3	( )	( )	opportunity	for	the party
4	( )	( )	opportunity	for	the government …
5	( )	( )	opportunity	for	us to think about …

Figure 1

Figure 2

Figure 3