

## 「論 文」

# A Meta-analysis of Data-driven Learning Approach in the Japanese EFL Classroom

Atsushi Mizumoto and Kiyomi Chujo

## Abstract

In this study, a meta-analysis was conducted targeting the studies that employ data-driven learning (DDL) approach in the Japanese EFL classroom context. After a thorough literature search, 32 effect sizes from 14 primary studies that took place in the Japanese EFL classroom were retrieved, coded, and calculated. The synthesized results, based on the classification of the outcome measures, showed that the DDL approach worked well particularly for learning vocabulary items (Level 1: lemma). It also worked positively for basic grammar items (Level 2: category) and noun and verb phrases (Level 3: phrase). For a proficiency measure, the combined effect size was small. Accordingly, the results of the current meta-analysis would provide further support for the use of DDL approach in the classroom, which could be an alternative methodology for facilitating the learning of lexico-grammatical items. Suggestions for further research and pedagogical implications are provided.

## 1. Introduction

The development of corpus linguistics as a discipline, especially since the end of the 20th century, has had a tremendous influence on the field of applied linguistics (Hunston, 2002). In particular, applied domains of corpus linguistics such as lexicography, pedagogic grammar (e.g., Biber, Johansson, Leech, Conrad, & Finegan, 1999), phraseology, and discourse analysis have benefited significantly from the very large corpora (Myles & Mitchell, 2004). Development of these applied domains within the field of corpus linguistics in turn has affected other areas in applied linguistics in general. Consequently, almost all introductory books on second language acquisition (SLA), language teaching, and language testing include sections on corpus linguistics or its applied domains (e.g., Loewen & Reinders, 2011; Long & Doughty, 2011; Mackey & Gass, 2012; Shohamy & Hornberger, 2008). For more specific pedagogical purposes, teaching and materi-

als development (Reppen, 2010; Tomlinson, 2013) have been advanced with the aid of corpus-based approaches. Without a doubt, the contribution of corpus application in applied linguistics is prevalently recognized as integral to the current development of the field.

As such, it seems reasonable for researchers and practitioners to try to make use of the potential applications of corpora in language teaching and learning (Aijmer, 2009; Aston, 2001; Flowerdew, 2012; O’Keeffe, McCarthy, & Carter, 2007; Sinclair, 2004). According to Römer (2010), pedagogical corpus applications are either direct (i.e., hands on for learners and teachers) or indirect (i.e., hands on for researchers and materials writers). Of these two types, direct applications of corpus, 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). Johns (1991) coined the term “DDL” more than 20 years ago, and DDL has been employed as a language learning methodology. The past decade has seen a growing body of research investigating the effects of DDL in the classroom (Tono, Satake, & Miura, 2014). Figure 1 shows the number of publications on DDL from 1989 to 2013, checked with the database, ProQuest (<http://search.proquest.com>). As can be seen from the figure, it is evident that the number of DDL studies has been on the rise, especially after 2000. In addition, research interest in DDL has been reflected in the fact that a special issue of *ReCALL* in 2014 was on DDL.

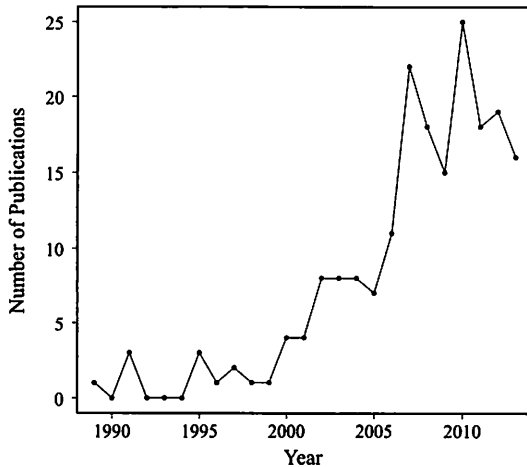


Figure 1: A ProQuest database search of DDL studies (1989–2013)  
Retrieved on November 8, 2014

With DDL receiving substantial attention from researchers and practitioners around the world, the common concern of stakeholders would be of course: “Is DDL effective as a teaching methodology?” Cheng (2010) maintained, “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” (p. 320). In this line of inquiry, Cobb and Boulton (2015) conducted a meta-analysis, which integrates the quantitative results gained from the past studies, and reported that, of the 21 studies out of 116 DDL studies (from 1989 to 2012) which met the requirements of meta-analysis, an effect size obtained for pre-post or within-group contrasts ( $k = 8$ ) was  $d = 1.68$ , 95% CI [1.36, 2.00], and for the between-group contrasts ( $k = 13$ ), the combined effect size was  $d = 1.04$ , 95% CI [0.83, 1.25]. They concluded that corpus use (DDL) in the classroom is effective on the ground that synthesized effect sizes of DDL studies are larger than those of meta-analyses of instructed SLA and CALL in general.

As the field of applied linguistics matures, more attention has been and will likely continue to be paid to research synthesis or meta-analysis (e.g., Norris & Ortega, 2000, 2006; Oswald & Plonsky, 2010). With meta-analysis, researchers can provide stronger evidence as it is an established method to integrate the results from the primary studies. The same trend is true for corpus linguistics, and more meta-analyses are appearing in the literature (e.g., Durrant, 2014; Jones & Kurjian, 2003).

From this perspective, the meta-analysis of DDL studies by Cobb and Boulton (2015) is certainly an invaluable piece of work and will contribute greatly to the understanding of how effective DDL is as a teaching approach. However, DDL studies in the context of Japanese EFL classrooms were not included in the meta-analysis by Cobb and Boulton (probably because some of the papers are written in Japanese). In this paper, therefore, with the aim of further shedding light on the effects of DDL as a teaching approach, we will focus on DDL studies in the Japanese EFL classrooms. We will conduct a synthetic investigation by means of meta-analysis and compare the result with that reported in Cobb and Boulton. The research question addressed in the current study therefore is:

How effective, in terms of synthesized effect size, is the DDL approach in the Japanese classroom context?

## 2. Method

### 2.1 Selecting Studies

An extensive literature search was conducted using databases such as ProQuest, which include those recommended by In'nami and Koizumi (2010) and Web of Science. In addition, academic search engines such as Google Scholar, Microsoft Academic Search, CiNii (Citation Information by NII), and J-STAGE (Japan Science and Technology Information Aggregator, Electronic) were used for identifying papers written in Japanese. The combination of following terms was used for the search for studies: (a) data-driven learning (or data driven learning), (b) corpus (or corpora), (c) concordance (or concordancer), (d) inductive, (e) Japan (or Japanese), and (f) English as a Foreign Language (EFL). In this way, all the relevant DDL studies conducted in the Japanese EFL context were retrieved and reviewed.

Because the purpose of the current meta-analysis was to investigate the effectiveness of a DDL approach in the Japanese EFL context, we included the studies that met the following eligibility criteria: (a) the study was conducted in Japan, (b) the study involved instruction with a DDL approach, (c) English was the target language in the class, and (d) tests were used as a quantitative measure of the effect of DDL. These criteria excluded some of the DDL papers written by researchers/practitioners based in Japan (Geluso & Yamaguchi, 2014; Hadley, 2002; Notohara, 2009; Quinn, 2013; Tono et al., 2014)

Through these screening procedures, 14 studies and 90 effect sizes from 656 participants, all of which were carried out by Chujo and her colleagues, were retrieved (Table 1). In the current meta-analysis, we focused on the pre-post or within-group contrasts because Cobb and Boulton's (2015) meta-analysis reported that synthesized effect size for between-group contrasts ( $k = 13$ ) was relatively large ( $d = 1.04$ , 95% CI [0.83, 1.25]), indicating the treatment group using a corpus in the classroom most likely always outperform the contrast group. In light of this meta-analytic result, we are now in a better position to investigate the effect of the DDL approach before and after the group receives the DDL instruction in the classroom (i.e., the pre-post or within-group contrasts). In addition, the result of the current meta-analysis, focusing on the pre-post or within-group contrasts, can be compared with that reported in other meta-analyses to further examine the relative strength of the effect sizes gained as a result of meta-analyses.

Table 1: Studies meta-analyzed in the current study

Study	<i>n</i>	Number of effect sizes	Number of procedures
1. Chujo & Oghigian (2007)	20	8	4
2. Chujo (2008)	75	4	2
3. Chujo et al. (2008)	21	2	1
4. Chujo, Anthony, & Oghigian (2009)	22	7	3
5. Nishigaki, Chujo, & Kijima (2010)	12	1	1
6. Chujo, Oghigian, & Nishigaki (2012)	15	3	1
7. Chujo et al. (2012)	66*	12	3
8. Chujo & Oghigian (2012)	39*	24	6
9. Nishigaki, Minegishi, & Chujo (2012)	27	1	1
10. Chujo, Oghigian, & Uchibori (2013)	50*	20	4
11. Nishigaki et al. (2013)	47*	2	2
12. Anthony et al. (2014)	103*	2	2
13. Chujo et al. (2014a)	145	1	1
14. Chujo et al. (2014b)	14	3	1
Total	656	90	32

*Note.* An asterisk (\*) indicates that the study included more than one group. Procedures include (a) Level 1 (lemma), (b) Level 2 (category), (c) Level 3 (phrase), and (d) Proficiency. Refer to Table 2 for details of the procedures.

## 2.2 Coding the Studies

The effect sizes from the collected 14 studies were coded and put into procedures depending on the test items in the primary studies (Table 2). First, test items were classified according to the measures (i.e., constructs intended to measure with those tests). The measures were then further divided into overarching procedures according to Pienemann's Processability Theory (1998). This is because, in the DDL practice in a series of studies conducted by Chujo and her colleagues, they employ a similar syllabus and classroom activities (e.g., use of concordancers and guided worksheets), based on the procedure level of Processability Theory.

**Table 2: Breakdown of procedures, measures, test items in the primary studies**

Procedures	Measures	Test items included	Number of effect sizes
Level 1 (lemma)	Vocabulary	Vocabulary items	6
	Grammar	Basic grammar items	6
Level 2 (category)	Category	(a) Word classes (b) Nouns (c) Adverbs (d) Derivations (e) Inflections	8
Level 3 (phrase)	NP structures	(a) Identifying NP (HFW) (b) Identifying NP (TOEIC words) (c) Producing NP (d) TOEIC-type NP	28
	VP structures	(a) Identifying VP (HFW) (b) Identifying VP (TOEIC words) (c) Producing VP (d) TOEIC-type VP	17
	TOEIC-type questions	TOEIC-type grammar items	16
Proficiency	TOEIC Bridge	TOEIC Bridge Test	9

*Note.* The procedure row corresponds with Piemann's Processability Theory (1998) except for proficiency. NP stands for noun phrases; VP stands for verb phrases. HFW is an abbreviation for high frequency words. "TOEIC-type questions" (Level 3) are complex and require learners to bring together knowledge of more than one aspect of grammar.

## 2.3 Analysis

Effect sizes (Cohen's  $d$ ) were calculated from means, standard deviations, and sample sizes. Because many studies did not report standard deviations, requests were sent to the authors of the original papers to obtain those missing values to compute the effect sizes. Effect size index  $d$ , for the between-group contrast, can be defined as the mean difference between two groups in standard deviation units. The formula is:

$$d = \frac{M_1 - M_2}{SD}$$

where  $M_1$  is the mean of one group (e.g., a treatment group);  $M_2$  the mean of the other group (e.g., a contrast group), and  $SD$  is the pooled standard deviation of the two groups (see Borenstein, Hedges, Higgins, & Rothstein, 2009 for the detail of the standardizer).

A complexity arises in the case of calculating  $d$  for the within-group or pre-post contrast (i.e., repeated measures), which can be defined as follows:

$$d = \frac{M_{post} - M_{pre}}{SD_{within}}$$

where  $M_{post}$  is the mean of the posttest,  $M_{pre}$  the pretest, and the standardizer,  $SD_{within}$ , is the standard deviation of the change (or gain/difference) scores, that is,  $M_{post}$  minus  $M_{pre}$  (Morris & DeShon, 2002). The standard deviation of the change scores however is not reported in the papers regularly; thus, meta-analysts cannot calculate the effect size  $d$  without contacting the author of the original article. Other researchers (Borenstein et al., 2009; Cooper, Hedges, & Valentine, 2009) recommend standardizers using the correlation between the pretest and the posttest. However, virtually no study reports the correlation between the pretest and the posttest, and researchers who meta-analyze within-group or pre-post contrast have to deal with the problem of missing correlation values between the pretest and the posttest scores. Thus, Cobb and Boulton (2015) used the formula to calculate  $d$  for the independent samples (i.e., between-groups contrast) and simply took the average of the effect sizes. Grgurović, Chapelle, and Shelley (2013) calculated the standardizer by averaging the standard deviations of pretest and posttest.

From a practical viewpoint, computing  $d$  for the independent samples or between-groups would not produce a meta-analytically synthesized result very much different from  $d$  for within-groups because they are identical when correlation between the pretest and the posttest is 0.5. According to Bonate (2000), “In psychological research the average correlation between measurements within an individual averages about 0.6” (p. 10). What meta-analysts need to do then is to report which formula they use for calculating the effect sizes and data used for the meta-analysis (i.e., means, standard deviations, sample sizes, and pre-post correlation when applicable) so that other researchers can reproduce and compare the results across meta-analyses.

Having been aware of the problem of missing correlation values in calculating the effect sizes for pre-post difference scores, we computed the effect sizes  $d$  with the following formula (Becker, 1988):

$$d = \frac{M_{post} - M_{pre}}{SD_{pre}}$$

where standardizer is the standard deviation of the pretest score. Bias correction factor was obtained using the following formula, and the bias correction fac-

tor was multiplied to obtain the corrected  $d$  (i.e.,  $d \times$  Bias Correction):

$$\text{Bias Correction} = 1 - \frac{3}{4(n-1) - 1}$$

where  $n$  is the sample size. In addition, to compute the sampling variance of effect size  $d$ , we used this formula (Kösters, Burlingame, Nachtigall, & Strauss, 2006):

$$v = \frac{2(1-r)}{n} + \frac{d^2}{2n}$$

As can be seen, for computing the bias correction factor and the sampling variance of effect size  $d$ , we needed the pre-post correlations for the studies, so we set all the values to 0.6 (i.e., the average correlation between measurements within an individual averages reported by Bonate, 2000). We also performed a sensitivity analysis by substituting a range of alternate pre-post correlations to make sure that the conclusions from the meta-analysis would not change drastically based on the imputed correlation values.

In meta-analysis, only one effect size from one study, representing a construct, should be retrieved to ensure statistical independence (Lipsey & Wilson, 2001). If more than one effect size (multiple effect sizes) come from the same sample, referred to as “stochastic dependency” (Plonsky, 2011), there are two ways recommended to reduce multiple effect sizes to a single effect size (Lipsey & Wilson, 2001, p. 113). The first one is to select only one effect size from multiple effect sizes. The second is to average several effect sizes to create a mean effect size for each construct. We applied the second method to our data and obtained the average effect size for each construct (i.e., Level 1: lemma, Level 2: category, Level 3: phrase, and Proficiency).

Weighting and averaging the effect sizes, as well as calculation of confidence intervals, were conducted with “metafor”: a meta-analysis package for R version 1.9.3 (Viechtbauer, 2010) of R version 3.1.1 (R Core Team, 2014). For the purpose of transparent sharing of data and results, all coded data and R codes used in this study are available online (<http://mizumot.com/files/ecs2015.html>). This will enable our readers to validate, scrutinize, and reproduce the results.

### 3. Results and Discussion

The result of the meta-analysis in the current study is presented in Figure 2. Overall effect size, which combines all the effect sizes with the random-effects



model, was 0.90, 95% CI [0.74, 1.07] ( $Q = 119.13$ ,  $df = 31$ ,  $p < .001$ ,  $I^2 = 80.16$ ). Without “Proficiency,” which is not the target of teaching with the DDL approach in the syllabus of Chujo and her colleagues, the synthesized effect size was 0.99, 95% CI [0.82, 1.17] ( $Q = 91.11$ ,  $df = 26$ ,  $p < .001$ ,  $I^2 = 79.10$ ). We present this result just as a reference because some of the effect sizes are from the same sample (i.e., stochastically dependent effect sizes).

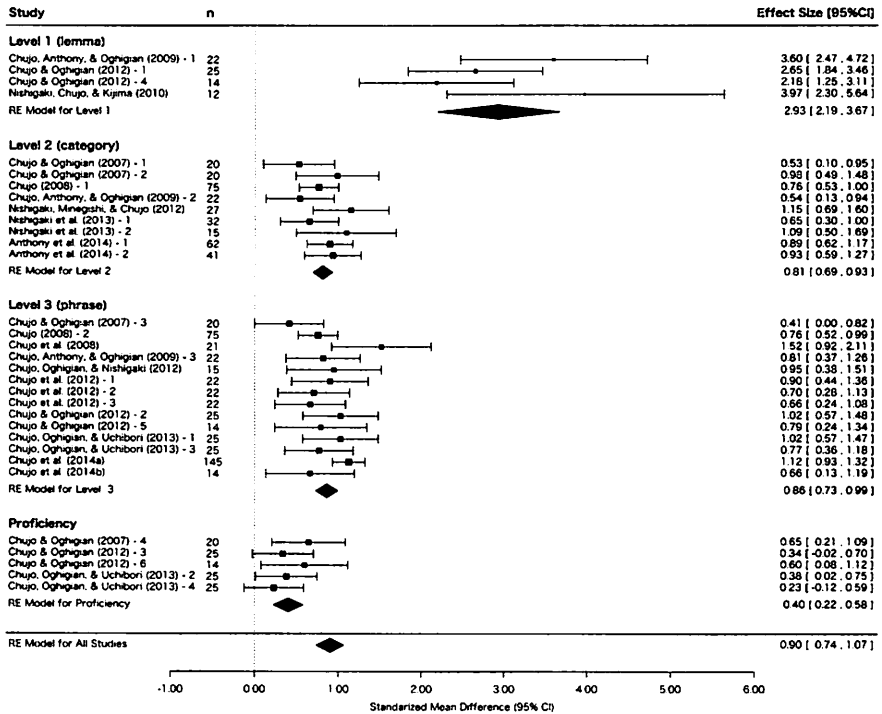


Figure 2: Results of the meta-analysis (random-effects model)

As expected, therefore, the  $Q$  value, the test for heterogeneity, and  $I^2$  show that the effect sizes are considerably different and heterogeneous across studies. Also publication bias (i.e., only publishing studies with statistically significant results) existed with the “funnel” and “regrest” functions in the metafor package in R at the overall level of meta-analysis. These results indicate that the current meta-analysis should be interpreted at each procedure level (i.e., Level 1: lemma, Level 2: category, Level 3: phrase, and Proficiency).

As for each level of procedures, the aggregated effect size for Level 1 (lemma) was 2.93, 95% CI [2.19, 3.67] ( $Q = 5.63$ ,  $df = 3$ ,  $p = .13$ ,  $I^2 = 47.21$ ); for Level 2 (category), 0.81, 95% CI [0.69, 0.93] ( $Q = 8.70$ ,  $df = 8$ ,  $p = .37$ ,  $I^2 = 0.01$ ); for Level 3 (phrase), 0.86, 95% CI [0.73, 0.99] ( $Q = 20.08$ ,  $df = 13$ ,  $p = .09$ ,  $I^2 = 34.69$ ); for Proficiency, 0.40, 95% CI [0.22, 0.58] ( $Q = 2.73$ ,  $df = 4$ ,  $p = .60$ ,  $I^2 = 0.00$ ). The  $Q$  value, the test for heterogeneity, and  $I^2$  indicate that the effect sizes across studies within each procedure (level) are homogeneous; thus, it proves that the meta-analysis in each procedure level was appropriate. Furthermore, publication bias was non-existent at the procedure levels.

Plonsky and Oswald (2014) proposed a field-specific benchmark of effect sizes for second language acquisition research based on 346 primary studies and 91 meta-analyses. According to their benchmark, for  $d$  values resulting from pre-post or within-group contrasts, a  $d$  value of .60 is generally considered small, 1.00 as medium, and 1.40 as large. Interpreting the magnitude of effect size according to Plonsky and Oswald's benchmarks, the synthesized effect size for all the studies was found to be of medium strength ( $d = 0.97$ ). Level 1 (lemma) showed the largest effect size of 2.93, although it should be noted that the number of studies included was small ( $k = 4$ ). The effect size of Level 2 (category) was 0.81 and that of Level 3 (phrase) was 0.86, which are close to medium strength. However, considering that the meta-analysis of pre-post design studies in CALL (Grgurović et al., 2013) reported the standardized mean gain of 0.35, 95% CI [0.26, 0.44] for the pre-post design studies ( $k = 16$ ), the synthesized effect sizes for Level 2 (category) and Level 3 (phrase) can be regarded as relatively large. The average effect size for Proficiency was small ( $d = 0.40$ ). This result is fully understandable given the fact that it takes considerable time and intensive training to record a sizable gain in proficiency tests; for example, it reportedly takes at least 100 hours of language training in the case of the TOEIC test (TOEIC Service International, 1999).

Cobb and Boulton (2015) reported a large effect size of 1.68, 95% CI [1.36, 2.00] for the pre-post or within-groups contrast ( $k = 8$ ) in their wider and more general DDL use in the classroom in other countries. Applying the same formulae used in our current meta-analysis (with the pre-post correlation set as 0.6), the effect size is still large:  $d = 1.53$ , 95% CI [0.85, 2.21] (although the homogeneity of the result is not confirmed,  $Q = 65.48$ ,  $df = 7$ ,  $p < .001$ ,  $I^2 = 91.58$ ). The positive results of the current meta-analysis in the Japanese EFL context, along with that reported in Cobb and Boulton, provide further evidence that the DDL approach can result in greater gains in the learning outcomes in Japanese EFL setting. Specifically, the DDL approach was more effective for Level 1 (lemma), which shows

the DDL approach is very promising for learning lexical items. Furthermore, for Level 2 (category) and Level 3 (phrase), substantial learning gains in terms of effect sizes were found as well. These findings should serve as supportive evidence that the DDL approach would be equally beneficial for learning basic grammar and formulaic sequences such as noun and verb phrases, as is intended by the syllabus, teaching procedures, and materials which have been used by Chujo and her colleagues for more than 10 years.

It should be pointed out that the current meta-analysis is limited in scope in that it included the primary studies all conducted by Chujo and her colleagues. Although we checked the publication bias, there is no means to defuse the researcher bias threat (i.e., researchers unconsciously influencing the result) to internal validity of the studies. It is therefore necessary for other researchers and practitioners (especially those in Japan) to conduct replication studies, the importance of which has been emphasized in the field of applied linguistics in recent years (Porte, 2012), with similar syllabi, teaching procedures, and materials employed by Chujo and her colleagues. Also, the proficiency of the participants in most of the studies in the current meta-analysis was rather low and homogeneous (approximately 300 to 350 in the TOEIC test); therefore, replication studies with higher proficiency learners would make a valuable contribution to teaching practice with the DDL approach in the Japanese EFL classroom.

In addition to these findings and possible limitations related to the current meta-analysis, we will address two crucial issues for facilitating a meta-analytic approach and improving the research quality in general. First, as is often pointed out in meta-analyses (e.g., Cobb & Boulton, 2015; Norris, 2012; Plonsky, 2011), the reporting practices of research papers need to be improved. Through our meta-analysis, we realized that essential information for conducting meta-analysis (such as descriptive statistic—especially standard deviations—and reliability indices of the measurement instruments) was frequently missing. As reproducibility is the core of scientific inquiry, researchers and practitioners in the field must adopt better reporting practices. Second, the terminology of DDL should be defined more clearly and precisely. In a series of DDL studies by Chujo and her colleagues, syllabus, all the teaching procedures, materials, and tested items are clearly defined so that readers of the paper can understand what they mean by DDL. However, without keeping in mind that their DDL approach is more like a supplementary aid for learning lexico-grammatical items with concordance lines, the term DDL may connote the original idea of “cutting out the middleman as much as possible” (Johns, 1991). In other words, researchers may have different ideas about DDL when they hear the term. In meta-analysis, as in all primary stud-

ies, defining constructs is of paramount importance because “analyses over disparate constructs are not generally meaningful” (Lipsey & Wilson, 2001, p. 113). This is also true of CALL research in general, but mixing different constructs (known as the “apples and oranges” problem) will lead to inconclusive and often contradictory results. Thus, careful examination of the construct is required when synthesizing research on DDL.

We have seen sizable gains in the learning outcomes in using the DDL approach in the classroom in this meta-analysis and in Cobb and Boulton (2015). Also, although Gilquin and Granger (2010) summarized the attitudes of learners toward DDL as “extremely mixed” (p. 365), in general, positive and favorable questionnaire responses toward the DDL approach have been reported from the learners in previous studies (e.g., Boulton, 2010; Chujo & Oghigian, 2012). Despite the accumulated evidence and on-going research interest in the DDL approach, not so many researchers and practitioners make use of its potential in their teaching practice. Gilquin and Granger argued that the problems and limitations of DDL are (a) the logistics, (b) the teacher’s point of view, (c) the learner’s point of view and (d) the content of DDL. In order for teachers to fully appreciate the benefits that DDL provides, advocates of the DDL approach will need to address these problems and limitations by supplying a complete concrete package, which includes an example of syllabi, teaching plans, sample lessons, materials, teaching manuals, and a user-friendly concordance tool (e.g., Chujo, Anthony, Utiyama, & Nishigaki, 2014).

#### 4. Conclusion

Ever since Johns (1991) proposed the notion of “data-driven learning,” DDL has attracted the attention of researchers and practitioners with the idea that it can be a tool to empower our students and make them autonomous in the long run. The field has accumulated a large body of knowledge about the use of DDL in the classroom, which has led to the current meta-analysis on the use of DDL in a specific EFL context in Japan.

In the current study, a meta-analysis has been conducted based on 32 effect sizes from 14 primary studies of DDL used in the Japanese EFL classroom context. The results were interpreted on the basis of the processing procedure levels (i.e., Level 1: lemma, Level 2: category, Level 3: phrase, and Proficiency. Level 1 (lemma) marked the largest effect size of standardized gain. Level 2 (category) and Level 3 (phrase) showed effect sizes of medium strength, which was larger than that gained from CALL research in general. For proficiency, the synthesized

effect size was small. These results corroborate findings by a similar meta-analysis by Cobb and Boulton (2015) targeting studies other than those conducted in Japan. Accordingly, the results of the current meta-analysis further evidence that the DDL works as intended and facilitates acquisition of lexico-grammatical items.

Of course, we do not aim to overgeneralize the results, nor do we go so far as to propose that DDL can replace conventional teaching practice in the EFL classrooms. Rather, the present meta-analysis confirms the positive benefits of applying a DDL approach to the learning of lexico-grammatical items. Using DDL in the classroom, as demonstrated in the primary studies of the current meta-analysis, will be a promising alternative, which practitioners could adopt with a degree of confidence in its potential to bring about positive and constructive change in the process of learning English. We hope a similar research endeavor will be organized to gauge the effectiveness of the DDL approach in the EFL settings.

### Acknowledgments

This study was supported by JSPS KAKENHI Grant Numbers 26704006 and 25284108. We would like to thank the anonymous reviewers for their constructive comments and feedback to improve the quality of the paper.

### References

References with an asterisk are those included in the meta-analysis.

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

\*Anthony, L., Chujo, K., Yokota, K., & Mizumoto, A. (2014, March). *Broadening the scope of parallel corpus tools: Using AntPConc in the DDL classroom*. Presentation given at Second Asia Pacific Corpus Linguistics Conference (APCLC) 2014, Hong Kong, China.

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

Becker, B. J. (1988). Synthesizing standardized mean-change measures. *Journal of Mathematical & Statistical Psychology*, 41, 257-278.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Harlow, England: Pearson Education.

Bonate, P. L. (2000). *Analysis of pretest-posttest designs*. Boca Raton, FL: Chapman & Hall/CRC Press.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex, England: John Wiley & Sons.

- Boulton, A. (2010). Data-driven learning: Taking the computer out of the equation. *Language Learning*, 60, 534-572. doi:10.1111/j.1467-9922.2010.00566.x
- Cheng, W. (2010). What can a corpus tell us about language teaching? In A. O’Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 319-332). London, England: Routledge.
- \*Chujo, K. (2008). Corpus-based syllabus design and its implementation. In J. Nakamura & S. Hotta (Eds.), *Intersection between corpus and English language education* (pp. 65-88). Tokyo, Japan: Shohakusha.
- \*Chujo, K., Anthony, L., & Oghigian, K. (2009). DDL for the EFL classroom: Effective uses of a Japanese-English parallel corpus and the development of a learner-friendly, online parallel concordancer. In M. Mahlberg, V. González-Díaz, & C. Smith (Eds.), *Proceedings of 5th Corpus Linguistics Conference 2009, University of Liverpool, UK*.
- \*Chujo, K., Anthony, L., Oghigian, K., & Uchibori, A. (2012). Paper-based, computer-based, and combined data-driven learning using a web-based concordancer. *Language Education in Asia*, 3, 132-145. Retrieved from [http://dx.doi.org/10.5746/LEiA/12/V3/I2/A02/Chujo\\_Anthony\\_Oghigian\\_Uchibori](http://dx.doi.org/10.5746/LEiA/12/V3/I2/A02/Chujo_Anthony_Oghigian_Uchibori)
- \*Chujo, K., Anthony, L., Utiyama, M., & Nishigaki, C. (2014b). Using the free-ware WebParaNews online concordancer in the EFL classroom. *Journal of the College of Industrial Technology, Nihon University*, 47, 49-63.
- \*Chujo, K., Mizumoto, A., Oghigian, K., Anthony, L., & Nishigaki, C. (2014a, September). Comparing DDL and non-DDL for different student learning styles. Poster presentation given at American Association for Corpus Linguistics (AACL) 2014, Flagstaff, Arizona.
- \*Chujo, K., Nishigaki, C., Uchibori, A., & Oghigian, K. (2008). An effective way to teach grammar to EFL beginners using data-driven learning. *Journal of the College of Industrial Technology, Nihon University*, 41, 15-33.
- \*Chujo, K., & Oghigian, K. (2007). Discovering grammar basics with parallel concordancing in the beginner-level EFL classroom. *Proceedings of The 2007 Asia TEFL International Conference, Kuala Lumpur, Malaysia*, 1-14.
- \*Chujo, K., & Oghigian, K. (2012). DDL for EFL beginners: A report on student gains and student views on the paper-based concordancing and the role of L1. In J. Thomas & A. Boulton (Eds.), *Input, Process and Product: Developments in Teaching and Language Corpora* (pp. 169-182). Brno, Czech Republic: Masaryk University Press.
- \*Chujo, K., Oghigian, K., & Nishigaki, C. (2012). Beginner level EFL DDL using a parallel web-based concordancer. *Proceedings of the FEELTA 2012, Far*

*Eastern Federal University, Vladivostok, Russia* (pp. 1-5).

- \*Chujo, K., Oghigian, K., & Uchibori, A. (2013). Comparing computer-based and paper-based DDL in the beginner level L2 classroom. *Journal of the College of Industrial Technology, Nihon University*, 46, 1-11.
- Cobb, T., & Boulton, A. (2015). Classroom applications of corpus analysis. In D. Biber & R. Reppen (Eds.), *The Cambridge handbook of corpus linguistics* (pp. 478-497). Cambridge University Press.
- Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis*. New York, NY: Russell Sage Foundation.
- Durrant, P. (2014). Corpus frequency and second language learners' knowledge of collocations. *International Journal of Corpus Linguistics*, 137, 443-477. doi:10.1075/ijcl.19.4.01dur
- Flowerdew, L. (2012). *Corpora and language education*. Basingstoke, England: Palgrave Macmillan.
- Geluso, J., & Yamaguchi, A. (2014). Discovering formulaic language through data-driven learning: Student attitudes and efficacy. *ReCALL*, 26, 225-242. doi:10.1017/S0958344014000044
- Gilquin, G., & Granger, S. (2010). How can DDL be used in language teaching? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 359-369). London, England: Routledge.
- Grgurović, M., Chapelle, C. A., & Shelley, M. C. (2013). A meta-analysis of effectiveness studies on computer technology-supported language learning. *ReCALL*, 25, 165-198. doi:10.1017/S0958344013000013
- Hadley, G. (2002). An introduction to data-driven learning. *RELC Journal*, 33, 99-124. doi:10.1177/003368820203300205
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.
- In'nami, Y., & Koizumi, R. (2010). Database selection guidelines for meta-analysis in applied linguistics. *TESOL Quarterly*, 44, 169-184. doi:10.5054/tq.2010.215253
- Johns, T. (1991). From print out to handout: Grammar and vocabulary teaching in the context of data-driven learning. In T. Johns & P. King (Eds.), *ELR Journal 4: Classroom Concordancing* (pp. 27-46). University of Birmingham.
- Jones, J. K., & Kurjian, J. (2003, October). *A meta-analysis of corpus-based research on morpheme frequency*. Paper Presented at Second Language Research Forum, Tucson, Arizona.
- Kösters, M., Burlingame, G. M., Nachtigall, C., & Strauss, B. (2006). A meta-analytic review of the effectiveness of inpatient group psychotherapy. *Group Dynamics: Theory, Research, and Practice*, 10, 146-163. doi:10.1037/1089-

2699.10.2.146

- Lipsey, M. W., & Wilson, B. D. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Loewen, S., & Reinders, H. (2011). *Key concepts in second language acquisition*. Basingstoke, England: Palgrave Macmillan.
- Long, M. H., & Doughty, C. J. (Eds.). (2011). *The handbook of language teaching*. Oxford, England: John Wiley & Sons.
- Mackey, A., & Gass, S. M. (Eds.). (2012). *Research methods in second language acquisition: A practical guide*. Chichester, England: John Wiley & Sons.
- Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, 7, 105-125. doi:10.1037//1082-989X.7.1.105
- Myles, F., & Mitchell, R. (2004). Using information technology to support empirical SLA research. *Journal of Applied Linguistics*, 1, 169-196.
- \*Nishigaki, C., Chujo, K., & Kijima, A. (2010). Teaching vocabulary to EFL advanced learners using data-driven learning with Japanese-English parallel-corpus. *Bulletin of the Faculty of Education, Chiba University*, 58, 279-286.
- \*Nishigaki, C., Minegishi, N., & Chujo, K. (2012). Data-driven learning in the junior and senior high school classroom. *Bulletin of the Faculty of Education, Chiba University*, 60, 419-426.
- \*Nishigaki, C., Yoshida, S., Kunikane, T., & Chujo, K. (2013, August). *Teaching vocabulary and grammar at junior high school by incorporating data-driven learning for facilitating focus on forms*. Presentation given at Language Education and Technology (LET) 2013, Tokyo, Japan.
- Norris, J. M. (2012). Meta-analysis. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 3653-3662). Malden, MA: Wiley-Blackwell. doi: 10.1002/9781405198431.wbeal0761
- Norris, M. J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417-528.
- Norris, M. J., & Ortega, L. (2006). *Synthesizing research on language learning and teaching*. Amsterdam, The Netherlands: John Benjamins.
- Notohara, Y. (2009). Pedagogical effects of convergent concordances: On low and high-intermediate adult Japanese EFL Learners. *Bulletin of the Graduate School of Education, Hiroshima University (Part. II)*, 58, 165-174. Retrieved from <http://ir.lib.hiroshima-u.ac.jp/ja/00028607>
- O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press.
- Oswald, F. L., & Plonsky, L. (2010). Meta-analysis in second language research:



- Choices and challenges. *Annual Review of Applied Linguistics*, 30, 85-110. doi:10.1017/S0267190510000115
- Pienemann, M. (1998). *Language processing and second language development: Processability theory*. Amsterdam/Philadelphia: John Benjamins.
- Plonsky, L. (2011). The effectiveness of second language strategy instruction: A meta-analysis. *Language Learning*, 61, 993-1038. doi:10.1111/j.1467-9922.2011.00663.x
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64, 878-912. doi:10.1111/lang.12079
- Porte, G. (Ed.). (2012). *Replication research in applied linguistics*. Cambridge University Press.
- Quinn, C. (2013). Students’ use of online reference tools for English composition and revision. *Departmental Bulletin Paper of Graduate School of Intercultural Studies, Kobe University*, 39, 19-42. Retrieved from [http://www.lib.kobe-u.ac.jp/handle\\_kernel/81005050](http://www.lib.kobe-u.ac.jp/handle_kernel/81005050)
- R Core Team. (2014). R: A language and environment for statistical computing (Version 3.1.1) [Computer software]. Vienna, Austria. Retrieved from <http://www.r-project.org/>
- Reppen, R. (2010). *Using corpora in the language classroom*. Cambridge University Press.
- Römer, U. (2010). Using general and specialized corpora in English language teaching: Past, present and future. In M. C. Campoy-Cubillo, B. Bellés Fortuño, & M. L. Gea-Valor (Eds.), *Corpus-based approaches to English language teaching* (pp. 18-38). London, England: Continuum.
- Shohamy, E., & 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.
- TOEIC Service International. (1999). *TOEIC User Guide*. Princeton, NJ: The Chauncey Group International.
- Tomlinson, B. (Ed.). (2013). *Developing materials for language teaching* (2nd ed.). London, England: Bloomsbury.
- Tono, Y., Satake, Y., & Miura, A. (2014). The effects of using corpora on revision tasks in L2 writing with coded error feedback. *ReCALL*, 26, 147-162. doi:10.1017/S095834401400007X
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1-48. Retrieved from

<http://www.jstatsoft.org/v36/i03/>

(Atsushi Mizumoto Kansai University E-mail: [mizumoto@kansai-u.ac.jp](mailto:mizumoto@kansai-u.ac.jp))

(Kiyomi Chujo Nihon University E-mail: [chujou.kiyomi@nihon-u.ac.jp](mailto:chujou.kiyomi@nihon-u.ac.jp))