

Using a Corpus in an EFL Classroom to Develop Depth of Vocabulary Knowledge

峰 松 和 子

MINEMATSU Kazuko

Abstract

A corpus is a collection of written or spoken texts, which is stored on a computer. This paper examines whether or not a corpus can be used to develop depth of vocabulary knowledge in an EFL classroom. The advantages of using a corpus lie in its authenticity and use of abundant examples. This suggests that a corpus could be used in fruitful ways to broaden or deepen learners' vocabulary. This paper focuses on the latter: developing depth of vocabulary knowledge. Data-driven learning (DDL) is one method of utilizing the corpus. First, this paper discusses the importance of developing depth of vocabulary knowledge. Then it investigates whether or not a corpus, especially DDL, can be used to attain this goal. Finally a pedagogical implication of utilizing DDL is suggested.

This paper makes clear that DDL may make it possible for learners to learn through concordancing, which would contribute to developing depth of vocabulary knowledge, including aspects such as grammatical functions, collocations and semantic prosody in an EFL classroom. It can be said that DDL is a promising approach in language teaching and learning in that it can expose learners to authentic language and motivate them to learn the target language.

1. Introduction

In recent years, there have been increasing attempts to apply a corpus to “fields as diverse as forensic linguistics and language teaching” (O’Keeffe, McCarthy, & Carter, 2007, p. 1). A corpus is a collection of written or spoken texts, which is stored on a computer. A corpus lets us look at a language feature in different ways. It is available for quantitative and qualitative analysis. As for quantitative analysis, for example, we can compare numbers of occurrences with frequencies in other corpora. As for qualitative analysis, we can see how a word or phrase is used across a corpus, which makes us look beyond the frequency of the word’s occurrence. In other words, it can help us “to see qualitative patterns of use beyond frequency” (O’Keeffe, McCarthy, & Carter, 2007, p. 3). This suggests that a corpus can be utilized in various ways in pedagogical contexts.

This paper examines whether or not a corpus can be used to develop depth of vocabulary

knowledge in an EFL classroom. In an EFL classroom, using a dictionary has been a norm when learners encounter unknown words except when they engage themselves in inferencing them. Both a dictionary and a corpus can be used for reference purpose. However, there is a difference between the use of a dictionary and that of a corpus. The advantage of using a corpus is that it shows many examples of the search item in its context of use. In addition, a corpus includes many examples of language which are actually used in authentic contexts. Its authenticity along with abundant examples suggest that a corpus could be used in fruitful ways to broaden and deepen learners' vocabulary in a pedagogical context. This paper focuses on the latter: developing the depth of vocabulary.

Vocabulary is very important in learning the second language (L2). In recent years, there is a tendency that vocabulary is no longer regarded as a single dimension but regarded as "a multidimensional construct" (Qian & Schedl, 2004, p. 29). According to Qian (2002), vocabulary knowledge comprises four intrinsically connected dimensions: vocabulary size, depth of vocabulary knowledge, lexical organization, and automaticity or receptive-productive knowledge. Past studies (Read, 1989; Wesche & Paribakht, 1996; Qian, 1999) show that vocabulary knowledge should at least comprise two dimensions, which are vocabulary breadth, or size, and depth, or quality, of vocabulary knowledge. The breadth of vocabulary has been mainly focused on learning L2 in EFL classrooms, but developing the depth of vocabulary knowledge should be emphasized more since it seems critical for learning the L2. The use of a corpus may well facilitate developing depth of vocabulary knowledge. It could help learners build an integrated lexicon and develop "learner agency" (Lantolf & Thorne, 2006). There may be various ways to do so by using a corpus and data-driven learning (DDL) can be one of the options. The organization of this paper is as follows: First, this paper discusses the importance of developing depth of vocabulary knowledge. Then it investigates whether or not a corpus, especially DDL can be used to develop depth of vocabulary. If so, how can it be used? Finally a pedagogical implication of DDL is suggested.

2. Importance of developing depth of vocabulary knowledge

According to O'Keeffe, McCarthy, and Carter (2007), the depth of knowledge is "the knowledge of the various aspects of use of a word, including, beyond its formal properties, its collocations, its sub-senses, and its semantic prosody" (p. 54). Haastrup and Henriksen (2000) show that depth of knowledge makes it possible for learners to create associations between words, and to place them meaningfully within various networks in relation to other words. Vocabulary breadth refers to "the number of words the meaning of which a learner has at least some superficial knowledge" (Qian & Schedl, 2012, p. 29). On the other hand, depth of vocabulary knowledge refers to "how well a learner know a word" (Qian & Schedl, 2012, p. 29). Pronunciation, spelling, meaning, register, frequency, and morphological, syntactic, and collocational properties should be covered in the depth dimension (Qian,

1999). The depth of vocabulary knowledge may help learners to create associations between words and to place them meaningfully within various networks in relation to other words (Haastrup & Henriksen, 2000). This suggests that “depth of knowledge is not simply a second-best to ever-increasing breadth” (O’Keeffe, McCarthy & Carter, 2007, p. 54). Qian (2002) shows the importance of vocabulary depth as well as vocabulary size in predicting performance on academic reading. Furthermore, Rashidi and Khosravi (2010) show that interrelations among depth of vocabulary knowledge, vocabulary size and reading comprehension are high and positive. In addition, they show that depth of vocabulary knowledge provided a significant contribution to the prediction of reading comprehension. Therefore, it is important to help learners to improve the quality or word knowledge, the depth of vocabulary knowledge.

3. Data-driven learning and the depth of vocabulary knowledge

3.1. Data-driven learning

Even though a communicative approach has emphasized the use of authentic texts, exercises in the course books or activities made by teachers in EFL classroom tend to be based on invented examples. Learners expose themselves to the language of course books, which may be different from actual language use, especially in the spoken language. DDL “brings authenticity into the classroom” (Gilquin & Granger, 2010, p. 359). By using corpora, learners can expose themselves to a large number of authentic instances of a particular linguistic item. This can contribute to “vocabulary expansion or heightened awareness of language patterns” (Gilquin & Granger, 2010, p. 59). Concordances can be produced in a number of formats (Kennedy, 1998), and the most usual form is the Key-Word-In Context (KWIC) concordance in which the collocates can be sorted based on words to the right or left of the key word. By looking at the center of the page, learners can focus on the main item of study.

The term data-driven learning (DDL) was coined by Johns (1988, 1991), who introduced concordancing in his work, *teaching English for Specific Purposes to non-native speakers of English*. According to Johns (1991), DDL is an inductive approach which relies on an “ability to see patterning in the target language and to form generalisations about language form and use” (p. 5). DDL can encourage noticing and consciousness-raising, which can “help [students] to become better language learners outside the classroom” (Johns, 1991, p. 31). DDL makes it possible for learners to get access to “more substantial amounts of corpus data than can be found in a dictionary, grammar or course book” (Chambers, 2010).

There are two ways to give learners access to corpus data: direct access by allowing learners to use corpora and concordancing softs, and indirect access by allowing them to learn about language use by studying concordances prepared in advance by the teacher (Chambers, 2010). Chujo, Anthony, Oghigian, and Uchibori (2012) show that, “with careful

guidance (i.e., a teacher-tested worksheet) and supervision, both computer-based and paper-based corpus exercises are equally effective” (p. 141). Furthermore, DDL may promote cognitive skills, many of which can apply to paper-based materials: “predicting, observing, noticing, thinking, reasoning, analysing, interpreting, reflecting, exploring, making inferences, focusing, guessing, comparing, differentiating, theorizing, hypothesizing, and verifying” (O’Sullivan, 2007, p. 277).

Boulton (2010) investigates how lower level learners cope with paper-based corpus materials and a DDL approach compared to more traditional teaching materials and practices. In his study, pretests and posttests show that “both are effective compared to control items, with the DDL items showing the greatest improvement, and questionnaire responses are more favorable to the DDL activities” (Boulton, 2010, p. 534). Even though empirical studies on the learning benefits of DDL are few, Cobb (1997) especially shows that using concordance line tasks led to vocabulary acquisition in his longitudinal study.

3.2. The use of DDL for developing depth of vocabulary

This section investigates whether or not DDL can be used to develop depth of vocabulary. According to Cobb (1997), DDL may have a role of providing multiple contextual encounters for the acquisition of new vocabulary, which is considered to be important for learning vocabulary. A corpus can be used to create a “‘lexico-grammatical profile’ of a word and its contexts of use” (O’keeffe, McCarthy, & Carter, 2007, p. 14). It means that a corpus can show typical contexts in terms of collocates, chunks/idioms, syntactic restrictions, semantic restrictions and prosody. Since a lexico-grammatical profile is drawn from concordance lines, it would give us the visible and productive patterning of a certain word or phrases. Kennedy and Miceli (2001) propose a four-stage search strategy for learners: At first, formulate the question. Then, devise a search strategy. Next, observe the examples and select relevant ones. Finally, draw conclusions. According to Chambers (2010), the essence of DDL is:

For learners to study the patterns of language use in a corpus, mostly through observing concordances, and work out for themselves how a word or a phrase is used. This process of inductive learning, in which the learners play an active part in the learning process is the essence of DDL. (p. 345)

This suggests that learners can discover frequent patterns by observing a large variety of examples of their use.

According to Nation (2001), aspects of word knowledge for testing include form, meaning and use. Use includes grammatical functions, collocations and constraints on use. As mentioned in Section 2, the depth of knowledge is the knowledge of the various aspects of use of a word, including its collocations and its semantic prosody. This paper mainly focuses on three aspects of depth of knowledge: grammatical functions, collocations, and semantic prosody.

3.2.1. Grammatical functions

At first, this section shows how DDL can be used for developing one aspect of vocabulary knowledge: grammatical functions of verbs. For example, one of the difficulties of the English language is that some verbs are followed by the gerund (e.g., doing) and others are followed by the infinitive (e.g., to do). Other verbs, however, can be followed by both. This subsection takes up these as examples, showing how DDL may facilitate learning grammatical functions.

[Three categories]

1) Verb+ gerund

Ex) admit doing, enjoy doing, avoid doing, finish doing, keep doing, deny doing, suggest doing, consider doing

2) Verb+ infinitive

Ex) agree to do, fail to do, hope to do, happen to do, learn to do, tend to do, decide to do, wish to do, pretend to do, want to do, refuse to do, hesitate to do, promise to do, manage to do, afford to do

3) Verbs that can be followed by both an infinitive and a gerund

Ex) start, begin, stop, remember

Learners look at concordance lines for these verbs to decide whether they belong to category 1, 2, or 3. Learners engage themselves in predicting, observing, noticing, thinking, reasoning, analysing, interpreting, reflecting, exploring and making inferences. Through this activity, the learners may nurture ability to see patterning and to form generalisations about language form and use, which may contribute to developing depth of vocabulary knowledge.

Next, this section focuses on empirical studies suggesting how DDL can be used for teaching grammar. Chujo and Oghigian (2008) show that a DDL approach can be effective in learning the structure of noun and verb phrases in the beginner level EFL classroom. They developed and tested a grammar-based syllabus which included a series of guided worksheets for pairwork-based tasks with a parallel Japanese-English corpus. Using a parallel corpora, they demonstrated that DDL can effectively be used with beginner-level classes. Furthermore, Chujo, Anthony, Oghigian, and Uchibori (2012) investigated the use of DDL in the beginner-level EFL classroom and compared the use of paper-based DDL, computer-based DDL, and combined DDL. They show that DDL is effective in teaching basic grammar such as noun phrases at the beginner level. It was also found that there are not any significant differences in effectiveness among the three approaches. In addition, students showed “significant gains with all three approaches, and all three clearly have distinct advantages” (Oghigian & Uchibori, 2012, p. 142).

3.2.2. Collocation

Collocation is defined as specified, identifiable, non-idiomatic, recurrent combinations (Benson, Benson & Ilson, 1997). Collocations are divided into two groups: grammatical and lexical collocations. The former consists of the main word plus a preposition or ‘to+ infinitive’ or ‘that-clause’. Grammatical collocation appears to be more difficult to learn than the lexical ones for the EFL learners (Benson et al., 1997). Especially, the usage of prepositions seems very difficult for learners of English. It is annoying for learners to decide or choose which preposition follows or precedes a certain word unless they have explicit knowledge about its collocation.

This subsection examines the past studies on whether the DDL approach is effective in learning collocation. Sun and Wang (2003) investigated the relative effectiveness of inductive and deductive approaches to learning collocations by using a concordancer. Participants were eighty-one senior high school students in Taiwan. It was found that the inductive group improved significantly more than the deductive group in the performance of collocation learning. In addition, it reveals that easy collocations seem to be more suitable in the concordancer learning setting. Furthermore, Jafarpour and Koosha (2013) showed that the DDL approach was highly effective in the teaching and learning of collocations of prepositions at three universities in Iran. It was also found that learners’ performance in collocation of prepositions was positively correlated with their levels of proficiency. They showed that the DDL instruction had an advantage over the conventional one. Students in the DDL groups outperformed those in the conventional groups. They suggest that “a greater explanatory power of the DDL” (p. 19) may contribute to highly significant results in the use of collocation of prepositions.

3.2.3. Semantic prosody

The term ‘semantic prosody’ was coined by Sinclair (1991) and Louw (1993). The concept of semantic prosody (Louw, 1993, p. 157) shows how certain forms can be imbued with “a consistent aura of meaning” by their collocates. Semantic prosody is “subtle aspects of meaning that are best revealed by the analysis of electronic text corpora” (p. 333). For example, concordances of *happen* show that things that happen are usually negative (Sinclair, 1987), and *utterly* is overwhelmingly combined with ‘bad’ adjectives (Louw, 1993). As other examples of semantic prosody, ‘end up’, ‘a recipe for’, and ‘signs of’ have negative connotations (Xiao & McEnery, 2006). As Louw (1993) suggests, relatively few semantic prosodies bear an “an affectively positive meaning” (p. 173). More than 90% of the collocates of *cause* are negative, for example *accident*, *cancer*, *commotion*, *crisis*, *delay* (Stubbs, 1995).

Yazdani and Darabi (2011) introduce a corpus-driven measure as a method to assess EFL learners’ knowledge of semantic prosody. They claim that “knowledge of semantic prosody is considered to be undermined by most EFL learners in receptive and, to a great extent, in productive mode” (p. 292). They consider that learning individual words and their meanings is not enough to achieve great fluency in a second language. Not only collocations but also

awareness of the conditions of semantic prosody is necessary. It means that “learners’ attention should be turned to these kinds of combination (words) and conditions (semantic prosody)” (Yazdani & Darabi, 2011, p. 292). They also suggest that teachers should realize the significance of semantic prosody in ESL/EFL learning and teaching.

3.3. Autonomous learner

In the DDL approach, learners are encouraged “to observe corpus data, make hypotheses and formulate rules in order to gain insights into language” (Gilquin & Granger, 2010, p. 359). It is a so-called inductive approach, and the learner is confronted with the data as directly as possible as if he or she were “a linguistic researcher” (Johns, 2002, p. 108). By including an element of discovery, DDL can make learning more motivating and more fun. DDL may provide a way “for students to take more active, reflective and autonomous roles in their learning” (Hyland, 2002, p. 129). DDL can “facilitate the development of learner autonomy” (Chamber, 2010), which has the effect of boosting his/her confidence and self-esteem. In that sense, it can be said that DDL can help learners to become a self-regulated learner. That perspective is especially important in learning vocabulary since “the vocabulary learning task is open-ended and impossible to complete in a typical institutional programme” (O’keeffe, McCarthy, Carter, 2007, p. 54). It is “learner agency” (Lantolf & Thorne, 2006) that makes the learner engage in active participation in the learning process.

4. Pedagogical implication

Through the use of DDL, “learners can acquire (or at least refine) a number of crucial learning skills” (Gilquin & Granger, 2010, p. 359) since this inductive approach basically involves cognitive processes (e.g., noticing, hypothesis formation, and hypothesis testing). This suggests that DDL may contribute to developing cognitive skills which “may also be transferred to other fields of study” (Gilquin & Granger, 2010, p. 359).

This paper also suggests that a DDL approach would bring about changes in language pedagogy in terms of the role of the teacher as well as the role of the learner. First, the role of the teacher differs fundamentally in that she or he is not a provider of the knowledge about the target language but a facilitator of the learning process. A teacher may play an important role in DDL, helping the learners to interpret the data, and giving them advice on how to search the corpus, which would lead them to analyzing their search results. In that sense, it requires teachers with new skills: knowledge about what learners can learn from a corpus, and how to analyze the data. Teachers also should know how to present the data of a corpus. According to Chambers (2010), “when faced with examples of variation in a corpus, a teacher is encouraged—or even obliged—to adopt a descriptive rather than a prescriptive approach to target language use” (p. 345). It means that consulting corpus data may change the relationship between the teacher and the target language.

5. Data-driven learning in an EFL classroom: Preparing and implementing DDL

This section deals with how a corpus can be used for data-driven learning (DDL) in an EFL classroom. At first, what a teacher should do in preparation for DDL is discussed. Then how DDL can be implemented in an EFL classroom is illustrated. Figure 1 shows what a teacher should do in preparation for DDL.

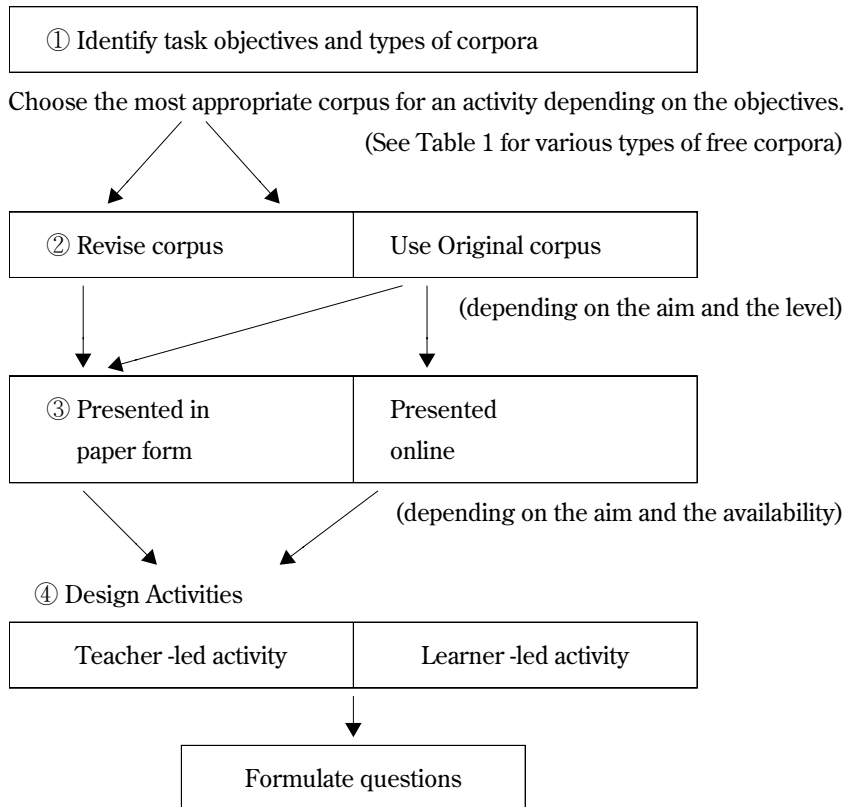


Figure 1 DDL. Selection Process

At first, a teacher should identify “task objectives and types of corpora” (Sripicharn, 2010, p. 372). The teacher chooses the most appropriate corpus for an activity depending on the objectives. Table 1 shows different types of free corpora which can be used in an EFL classroom.

After identifying task objectives and types of corpora, a teacher decides whether the corpus should be revised or the original corpus can be used for learners. Depending on the aim of the activity and the level of learners, the concordance lines may be edited or presented in their original form. Boulton (2009 b, p. 89) is not in favor of manipulation such as editing the concordances, simplifying them, since it would undermine the ‘authenticity’ advantage of

Table 1 Free corpora**A. General corpora**

British National Corpus (BNC) < http://www.natcorp.ox.ac.uk/ >	The BNC is a 100 million word collection of samples of written and spoken language from a wide range of sources.
Corpus of Contemporary American English (COCA): < http://www.americancorpus.org/ >	COCA has more than 450 million words of spoken and written American English.

B. Some specialized corpora

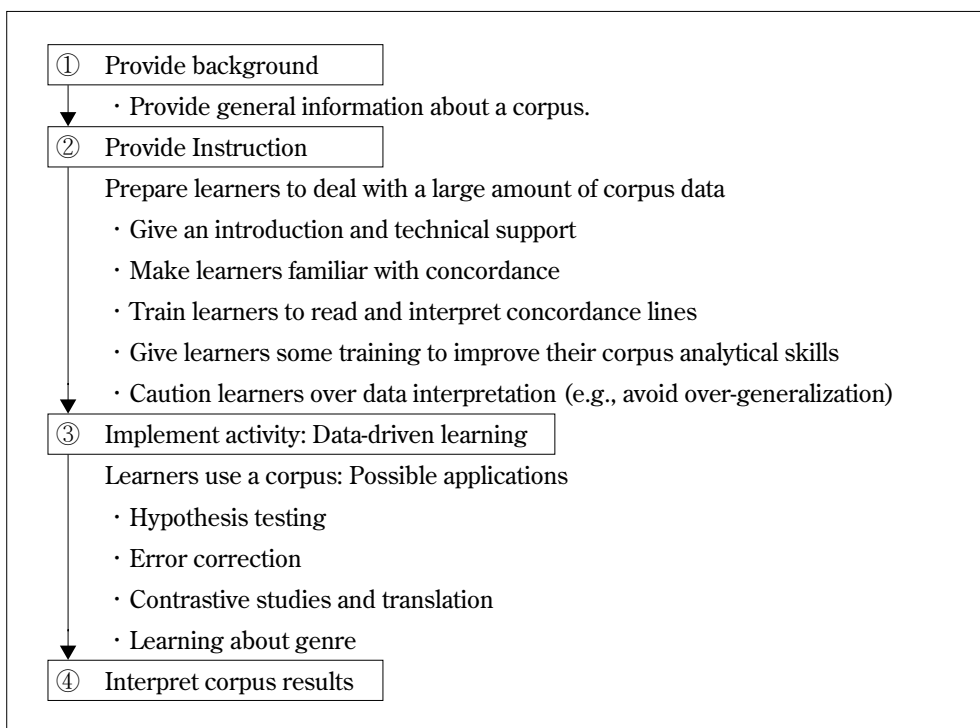
Michigan corpus of Academic Spoken English (MICASE): < http://quod.lib.umich.edu/m/micase/ >	MICASE has approximately 1.8 million words of spoken academic American English. Searches based on speaker attributes (age; gender; native speaker status; etc.) are possible.
Business Letter Corpus (BLC): < http://www.someya-net.com/concordancer/index.html >	BLC is made up of three sub-corpora of business letters, memos and reports. These are authentic business texts (over 1 million words in size).
Web Para News: < http://www.antlab.sci.waseda.ac.jp/webparanews/ >	Web Para News is a parallel corpus of English and Japanese, which uses news articles from a database of the National Institute of Information and Communications Technology (NICT).

DDL. However, according to Gilquin and Granger (2010), “manipulation is sometimes necessary in DDL (especially with beginners)” (p. 363).

Next, concordances may be presented on screen or in a handout depending on learners’ level as well as the availability of hardware and software. Once determined, activities can be either learner-led activities or teacher-led activities. For learner-led activities, “discovery learning is often claimed to be most appropriate for ‘very advanced learners who are filling in gaps in their knowledge rather than laying down the foundations’” (Gilquin & Granger, 2010, p. 363), whereas, “teacher-led activities tend to be better suited to beginners” (p. 363). Finally, a teacher designs activities for DDL. In DDL, learners use language corpora “to uncover language patterns and use” (Gilquin & Granger, 2010, p. 372), taking the role of researcher or language detective.

Next, Table 2 illustrates how a corpus can be used for DDL in an EFL classroom (Sripicharn, 2010). Before using a corpus in an EFL classroom, it is necessary for the teacher to find out what students know about the corpus and provide general information. For example, “the nature of corpus data, different types of corpora, basic functions of concordancing tools, the unique characteristics of concordance lines, and strengths and

Table 2 How a Corpus can be Used for DDL in an EFL classroom



limitations of language corpora for the purpose of language analysis” (Sripicharn, 2010, p. 372). In order to prepare learners to deal with a large amount of corpus data, “teachers may use a scaffolding technique to help students become familiar with the data, starting with handouts with short tasks using clear examples for the corpus data” (Sripicharn, 2010, p. 375). A corpus can be used in several ways for DDL, for example, for hypothesis testing or

Table 3 Areas of corpus investigation and sample questions

Areas of investigation	Sample questions
Grammatical questions	<ul style="list-style-type: none"> · Is ‘advice’ a count or non-count noun? · What is the noun suffix of ‘employ’?
Collocation/ Phraseological	<ul style="list-style-type: none"> · What common adjectives used in front of ‘argument’? · What words can we use to describe our body parts (e.g., ‘hair, nose’, ‘eyes’)?
Connotation/ semantic	<ul style="list-style-type: none"> · What are the differences between ‘childish’ and ‘childlike’?
prosody	<ul style="list-style-type: none"> · What are the differences between ‘utterly’, ‘bittery’ and ‘absolutely’?

(Source: Sripicharn, 2010, P. 377)

for error correction. For hypothesis testing, learners observe corpus data, make hypotheses and formulate rules in order to gain insights into language (Sripicharn, 2010). For error correction, the students can use corpus data to revise their first-draft writings in a number of areas of grammatical and lexical errors (O'Sullivan & Chambers, 2006). After observing, analysing, and making inferences, learners may interpret corpus results working in pairs or in groups.

A corpus can be used to develop depth of vocabulary knowledge such as grammatical forms, collocation and semantic prosody. Areas of corpus investigation and sample questions designed by Sripicharn (2010, p. 377) are shown in Table 3.

Next, Table 4 shows a set of teacher-selected concordance lines and questions designed to raise awareness of semantic prosody.

Table 4 an Example of DDL (semantic prosody)

(Click **keyword** for Larger Context)

the Social Charter would “**CAUSE** a great increase in unemployment not only which disrupt peace **and CAUSE** death and destruction, Mr Hrawi said. create confusion **and CAUSE** distress. Changes of staff during the of flexibility which **can CAUSE** problems during installation. Fibreglas indexing of keywords **can CAUSE** much confusion for the searcher.

(ESL Teaching & Learning Resources: Online Concordancers

available online <http://www.lex Tutor.ca/concordancers/>)

Questions)

1. Does the word 'cause' collocate with positive words or negative words?
2. Write the words which come after the word 'cause', representing either a positive or a negative situation. (e.g., 'great increase in unemployment')

6. Problems and limitations of DDL

According to Gilquin and Granger (2010), there are four aspects of DDL that may be problematic: “the logistics, the teacher’s point of view, the learners’ point of view and the content of DDL” (p. 366). First, learners need not only computers but also corpora and text retrieval software. Therefore, teachers should create their own materials, which might be time-consuming. Secondly, the teacher may have little knowledge about how to use corpora in the classroom. It may be necessary for teachers to have in-service teacher training programs about corpora. Thirdly, learners need to have some training to acquire the basic skills which can be called corpus literacy. Without it, it is possible for the students to draw wrong inferences in DDL. It implies that DDL might be “suitable for certain learners only,

depending on their learning style” (Gilquin & Granger, 2010, p. 367). Finally, as for the content of DDL, there are some problems; (a) the corpus could show more details than the student is expected to learn, (b) the corpus could contain the language such as swear words or non-standard forms, and (c) it may not be effective for all aspects of language.

Furthermore, Chujo, Anthony, Oghigian, and Uchibori, (2010) point out that “students are often more comfortable *studying* about a language rather than *producing* a language” (p. 142). They suggest that it is important to include production so that students can apply what they are learning to written or spoken constructions.

7. Conclusion

This paper makes clear that DDL may make it possible for learners to learn through concordancing, which would contribute to developing the depth of vocabulary knowledge: grammatical functions, collocations and semantic prosody in an EFL classroom. However, DDL has some problematic aspects and many more empirical studies are necessary to demonstrate the value of DDL in an EFL classroom. Nevertheless, it can be said that DDL is a promising approach in language teaching and learning in that it can expose learners to authentic language and motivate them to learn the target language, engaging themselves in discovery.

References

- Ahmadian, M., Yazdani, H., & Darabi, A. (2011). Assessing English learners' knowledge of semantic prosody through a corpus-driven design of semantic prosody test. *English language Teaching*, 40 (4), 288-298.
- Benson, M., Benson, E, & Ilson, R. (1997) [1986] The BBI dictionary of English word combinations (2nd edition). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Boulton, A. (2010). Data-Driven Learning: Taking the computer out of the equation. *Language Learning*, 60 (3), 534-572.
- Chambers, A. (2010). What is data-driven learning? In O'Keeffe, A & MacCarthy, M. (Eds.), *The Routledge handbook of corpus linguistics* (pp. 345-358). London: Routledge.
- Chujo, K., & Oghigian, K. (2008). A DDL approach to learning noun and verb phrases in the beginner level EFL classroom. In A. Frankenberg-Garcia, I. Rkibi, M. R. Cruz, R. Carvalho, C. Direito, & D. Santons-rosa (Eds.), *Proceedings of TaLC 8-Lisbon, 8th Teaching and language corpora conference* (pp. 65-71). Lisbon, Portugal: Associacao de Estudos e de investigacao Cientifica do ISLA-Lisbo Fichier.
- Chujo, K., Anthony, L., Oghigian, K., & Uchibori, A. (2012). Paper-based, computer-based,

- and combined data-driven learning using a web-based concordance, *Language Education in Asia*, 2012, 3 (2), 132-145.
- Cobb, T. (1997). Is there any measurable learning from hands-on concordancing? *System* 25 (3): 301-15
- Cobb, T. & Greaves, C. (2012). ESL Teaching & Learning Resources: Online Concordancers. Retrieved from <http://www.lex tutor.ca/concordancers/>
- Gilquin, G. & Granger, S. (2010). How can DDL be used in language teaching? In O'keeffe, A., & McCarthy, M. (Eds.). *The Routledge handbook of corpus linguistics* (pp. 359-369). London: Routledge.
- Haastrup, K. & Henriksen, B. (2000). Vocabulary acquisition: Acquiring depth of knowledge through network building. *International Journal of Applied Linguistics*, 10 (2), 221-240.
- Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition* 21, 303-317.
- Horst, M. & Cobb, T. (2006). Editorial: Second Language vocabulary acquisition. In Horst, M. & Cobb, T. (Eds.), Second special vocabulary edition of *Canadian Modern Language Review*, 63 (1), 1-12.
- Hyland, K. (2002). *Teaching and researching writing*. London: Pearson Education.
- Jafarpour, A. & Koosha, M. (2013). Data-driven learning and teaching collocation of prepositions: The case of Iranian EFL students. *Research on Foreign Languages Journal of Faculty of Letters and Humanities*, Year 49, No. 200. 1-31.
- Johns, T. (1988). Whence and whither classroom concordancing? In Bongaerts, T., De Haan, P., Lobbe, S. & Wekker, H. (Eds), *Computer applications in language teaching* (pp. 9-27). Dordrecht: Foris.
- Johns, T. (1991). Should you be persuaded: Two Examples of Data-driven learning. In T. F. Johns & P. King (Eds). *ELR Journal Vol.4: Classroom concordancing* (pp. 1-6). Birmingham: CELS, University of Birmingham, 1-6.
- Johns, T. (2002). Data-driven learning: the perpetual challenge. In Kettemann, B. & Marko, G. (Eds.), *Teaching and learning by doing corpus linguistics* (pp. 107-117). Amsterdam: Rodopi,
- Kennedy, G. D. (1998). *An introduction to corpus linguistics*. London; New York: Longman.
- Kennedy, C. & Miceli, T. (2001). An evaluation of intermediate students' approaches to corpus investigation, *Language Learning and Technology* 5 (3), 77-90
- Lantolf, J. P. & Thorne, S. (2006). *Sociocultural theory and the genesis of second language development*. Oxford: Oxford University Press.
- Louw, W. E. (1993). "Irony in the Text or Insincerity in the Writer? The Diagnostic Potential of Semantic Prosodies". *Text and Technology: In Honour of John Sinclair*. 157-176. Amsterdam: John Benjamins.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and*

- teaching*. Cambridge: Cambridge University Press.
- Oster, U. & Lawick, H. (2008). Semantic preference and semantic prosody: A corpus-based analysis of translation-relevant aspects of the meaning of phraseological units. *Translation and meaning, Part 8*, 333-344
- O'Sullivan, I. (2007). Enhancing a process-oriented approach to literacy and language learning: The role of corpus consultation literacy. *Cambridge Journals Online-ReCALL, Vol. 19* (03)
- Qian, D. (1999). Assessing the role of depth and breadth of vocabulary knowledge in reading comprehension. *The Canadian Modern Language Review, 56*, 282-308.
- Qian, D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: an assessment perspective. *Language Learning, 52*, 513-36.
- Qian, D. & Schedl, M. (2004). Evaluation of an in-depth vocabulary knowledge measure for assessing reading performance. *Language Testing, 21* (28), 28-52.
- Rashid, N., & Khosravi N. (2010). Assessing the role of depth and breadth of vocabulary knowledge in reading comprehension of Iranian EFL learners. *Pan-Pacific Association of Applied Linguistics 14* (1), 81-108.
- Read, J. (1989). Towards a deeper assessment of vocabulary knowledge. Paper presented at the 8th World Congress of Applied Linguistics. Sydney, NSW, Australis, August, 1987. ED 301048. Washington, DC: ERIC Clearinghouse on Languages and Linguistics.
- Sinclair, J. (1987). The nature of the evidence. In Sinclair, J. McH. (Ed.), *Looking up*. (pp. 150-159). Glasgow: Collins.
- Sinclair, J. (1991). "Words and phrases". *Corpus, Concordance, Collocation* (pp. 70-75). Oxford: Oxford University Press.
- Sripicharn, P. (2010). How can we prepare learners for using language corpora? In O'keeffe, A., & McCarthy, M. (Eds.). *The Routledge handbook of corpus linguistics* (pp. 371-384).
- Stubbs, M. (1995). Corpus evidence for norms of lexical collocation. In Cook, G. & Seidhofer, B. (Eds.), *Principle and practice in applied linguistics: Studies in Honour of H.G. Widdowson* (pp. 245-256). Oxford: Oxford University Press.
- Sun, Y. and Wang, L. (2003). Concordancers in the EFL Classroom: Cognitive approaches and collocation difficulty. *Computer Assisted Language Learning, Vol. 16* (1), 83-94.
- Wesche, M. & Paribakht, T. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. *Canadian Modern Language Review, 53*, 13-40.
- Xiao, R. & McEnery, T. (2006). Collocation, semantic prosody and near synonymy: A cross-linguistic perspective, *Applied Linguistics 27*(1), 103-29.
- Yazdani, H. & Darabi, A. (2011). Assessing English learners' knowledge of semantic prosody through a Corpus-Driven Design of semantic prosody test. *English Language Teaching, 4* (4), p. 288.
- Yoon, H. & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing, 13*, 257-284.