



# Using Corpus-Aided Data-Driven Learning to Improve Chinese EFL Learners' Analytical Reading Ability

Manfei Xu<sup>(✉)</sup>, Xiao Chen, Xiaobin Liu<sup>(✉)</sup>, Xiaoyue Lin,  
and Qiaoxin Zhou

School of Foreign Studies, South China Normal University,  
55 Zhongshan West Ave, Shipai, Tianhe, Guangzhou, China  
{xumanfei, liuxiaobin}@m.scnu.edu.cn

**Abstract.** Data-driven learning (DDL) has been proved successful in improving learners' English language skills (Chen and Flowerdew 2018; Larsen-Walker 2017; Gui et al. 2010; He 2010). This study investigates a relatively under researched aspect, i.e., whether corpus-aided DDL can help learners improve their analytical reading ability. In a 6-week teaching experiment with a class of 31 advanced Chinese EFL learners, different corpus processed materials such as *keywordlists* and *concordances* were used to assist the subjects towards a deeper and quicker understanding of the texts. A post-study questionnaire survey and interviews obtained positive comments by the subjects. Based on the students' feedback, a reading-writing experiment was further conducted to three more classes (135 students in total) using three different instructional modes, i.e., sheer teacher instruction, sheer corpus processed input, and teacher instruction plus corpus processed input. The study found corpus-aided DDL targeted at learner needs effective in improving the subjects' reading capacity.

**Keywords:** Corpus-aided input · DDL · Analytical reading

## 1 Introduction

Integration of technology with education has become an ideology of teaching and learning highly sought after for the last two decades as it brought about new opportunities, resources, useful tools and methods to teachers and learners. With regard to language teaching and learning, data-driven learning (DDL) advocates learner engagement through technology means which has been found effective in enriching EFL learners' learning experience and providing opportunities for them to explore language knowledge and language use. Researchers reported on how corpus-based or corpus-aided DDL has been used in assisting learners with their English language development (Chen and Flowerdew 2018; Larsen-Walker 2017; Gui et al. 2010; He 2010).

While the DDL approach has been applied to learners' English writing, grammar or vocabulary enhancement, a relatively less researched on aspect concerns the influence of DDL on the development of advanced EFL learners' analytical reading ability.

The current study designed a 6-week teaching using different corpus processed materials such as *wordlists*, *keywordlists*, and *concordances* to a class of 31 advanced Chinese EFL learners with a purpose of helping them to gain a deeper and quicker understanding of the texts. A post-study survey was conducted which found that the subjects responded to the approach positively. In order to verify some feedback by the subjects with regard to the corpus-aided approach, a reading-writing experiment was further conducted to three more parallel classes (135 students in total) using three different instructional modes. Class A were required to write a summary after reading a text with teacher instruction, while Class B finished the writing after reading the same text but with additional corpus-aided input. Class C finished the same task with both teacher instruction and corpus-aided input. The data obtained provide further evidence that DDL is effective in helping the subjects establish a deeper understanding of the text, and that the students performed the best under the mode of corpus-aided input with teacher instruction.

## 2 Corpus-Aided Reading Teaching for EFL Learners

Data-driven learning is a special language teaching and learning approach which enables students to make use of corpus to explore specific linguistic issues through authentic language materials. The theory underpinning it is corpus linguistics, which could be applied far beyond the discipline itself (McEnery et al. 2006: 8). Cobb and Boulton (2015) suggest that its methodologies make great influence on improving description of language varieties and features, which are important factors in language teaching. The activities design of DDL, namely, corpus-based learning, covers a wide range from “hard” to “soft” (Bernardini 2004; Gabrielatos 2005), depending on the degree learners themselves can take over the learning stages (Cobb and Boulton 2015).

For its diverse advantages, DDL is a frequently discussed issue in language teaching. Cobb and Boulton (2017) review 64 studies and find that most of the relevant studies focus on corpus-based learning of vocabulary from the lexicogrammar or grammar perspective. They argue DDL plays an excellent role in helping learners deal with the great difficulties they encountered in producing natural and effective language. Johns (1991) analyses two samples of DDL materials and states that this approach provides a new style of “grammatical consciousness-raising” (Rutherford 2014), which is different from the traditional teaching approach that presents students with a set of “rules” directly.

Apart from the studies which focus on grammar and vocabulary, few scholars pay attention to the application of corpus to reading teaching. The feasibility and advantages were not fully researched on. Cobb and Boulton (2015) point out that corpus tools can be flexibly adjusted and applied to individual texts. It is helpful in deciding what elements of a text can be paid great attention to. Such concentration can help teachers decide on reading activities to be conducted. Except for its feasibility, research shows the strengths of incorporating technology with reading instruction. After reviewing 42 studies, scholars (Jamshidifarsania et al. 2019) agree that the combination of technology and reading instruction is able to increase motivation, reduce cognitive load and allow students learn at their own pace. A higher motivation coupled with less reading

anxiety is useful in improving reading effects. Others (Kamil and Chou 2009; Boekaerts and Corno 2005; Zimmerman and Tsikalas 2005) point out the potential for teaching reading through computer-based approach and the benefits it holds. As one of the pioneers, Hadley and Charles' (2017) study reveals the effects of DDL in reading instruction and confirms that this method has positive effect on the extensive reading program the study designed.

Studies cited above laid some foundation for the corpus-aided approach in reading teaching the present study attempts to explore. What sets it apart from previous studies is its concentration on enhancing students' analytical reading which is a kind of in-depth reading. According to Chall (1983) who divides reading ability development into five stages based on Piaget's cognitive development theory, college students are at the stage of having their analytical reading ability quickly developed. Learners at such a stage attempt to establish their view of the world via reading and analyzing different types of texts. They make inferences, judgments, and draw their own conclusions by interacting with the texts. Besides, they develop their logical thinking abilities by analyzing the themes contained in the texts, textual development features, language and rhetoric features, etc. For English major students at college level, they are required to develop excellent English skills and are expected to enhance their critical thinking ability which can be demonstrated through analytical reading performance. Such an expectation is clearly indicated in the nationally used *Syllabus of English Teaching for English Majors of Higher Education* published by China's Ministry of Education (2000).

As improving Chinese advanced EFL learners' analytical reading ability becomes a goal in English education currently, a question concerning how it can be empirically done arise. Besides, how the teaching can be assisted through technology to improve its effects is also to be considered. The corpus-aided approach of the study makes use of a user-friendly corpus tool, AntConc 3.2.0. (Anthony 2014) which can quickly present very useful information of a text or texts such as *keywordlist*, *concordances* and *n-grams* (clusters of different number of words centering around certain words). With this core information, teachers can guide students towards understanding of the theme of a text, considering how it develops into sub-themes, and what and how language forms are used for the making of aboutness of a text. Besides such guided analysis which students can do, their reading effects can be enhanced because DDL approach provides them with autonomous chances of exploration, i.e., they are engaged in making attemptive analysis of different kinds of materials the corpus tools instantaneously offer.

### 3 Research Design

#### 3.1 Research Questions

In order to examine the effectiveness of corpus-aided teaching of reading, the study aims at answering the following questions: (1) How can the teaching of analytical reading make use of the corpus method and tools? (2) Can corpus-aided DDL help Chinese advanced EFL learners improve analytical reading performance?

To answer the first question, a 6-week teaching experiment of corpus-aided instruction was conducted. To answer the second question, a post-study questionnaire survey, interviews, and a reading-writing experiment were carried out.

### 3.2 Participants

The teaching experiment with the questionnaire survey and interviews took place in April to May of 2016. A class of 31 students participated. They were third-year English majors who had all passed *Test for English Major Band 4* with an English proficiency level comparable to IELTS 6.0–6.5. They had taken the course of Corpus Linguistics in the previous term and were familiar with the corpus tools used in this study. In the second stage of the study, three classes of 135 second-year English majors participated in a 40-min reading-writing experiment. They had also taken the course of Corpus linguistics and were familiar with corpus tools used in this study.

### 3.3 The Corpus and Tools

A mini corpus of five reading passages taken from an intensive reading course for English major students was constructed for the teaching experiment. Table 1 shows the titles and lengths of the texts.

**Table 1.** The five texts used in the study

Title	The science of custom	Beauty	Appetite	Euphemism	Teaching as mountaineering
Word count	876	1007	733	1332	1578

The corpus software used in the study is AntConc (Anthony 2014). The main tools used are *keywordlist* and *concordance*. *Keywords* in Corpus linguistics refers to a list of words of a text (or texts/corpus) which are unusually frequent (with statistic significance) when compared with a wordlist of a referent text (or texts/corpus). The reference is of data from different topics and of a larger word size (at least five times more). Figure 1 shows a part of the keywordlist extracted from the text *The Science of Custom*. The higher the keyness value, the greater the difference in word frequency between the target text and the reference corpus. Keyness is regarded as a mostly textual quality (Scott and Tribble 2006). Keywords often indicate the aboutness and stylistic features of a text.

Types Before Cut: 357			Types After Cut: 320	Types Before Cut: 357			Types After Cut: 320
Rank	Freq	Keyness	Keyword	Rank	Freq	Keyness	Keyword
1	9	62.562	custom	17	3	20.854	tradition
2	8	55.611	study	18	3	20.854	traditional
3	7	48.659	social	19	3	16.418	ourselves
4	24	48.448	is	20	3	16.418	possible
5	52	46.426	of	21	5	16.290	those
6	9	39.557	our	22	8	15.177	any
7	5	34.757	customs	23	3	14.250	part
8	7	26.253	own	24	2	13.903	anthropologist
9	15	22.385	his	25	2	13.903	behavior
10	3	20.854	anthropology	26	2	13.903	belief
11	3	20.854	behaviour	27	2	13.903	beliefs
12	3	20.854	born	28	2	13.903	civilization
13	3	20.854	habits	29	2	13.903	community
14	3	20.854	human	30	2	13.903	cultures
15	3	20.854	individual	31	2	13.903	experience
16	3	20.854	necessary	32	2	13.903	forms

Fig. 1. Extract of keywordlist

13 ill have reference to his particular traditional customs. John Dev  
 14 of thinking. Even in his philosophical probings he cannot go behi  
 15 nd these stereotypes; his very concepts of the true and the false  
 16 l relation to that of our civilization. To the anthropologist, our  
 17 o the anthropologist, our customs and those of a New Guinea tribe  
 18 n belief over against our neighbour's superstition. It was necess  
 19 considered together, our own among the rest.  
 20 here we no longer set our own belief over against our neighbour's  
 21 The inner workings of our own brains we feel to be uniquely worthy  
 22 ped by one tradition, our own, but as it has been shaped by any t  
 23 other societies than our own. For its purposes any social regulat  
 24 is as significant as our own, though it may be that of the Sea Dy  
 25 distinctions between ourselves and the primitive, ourselves and t  
 26 es and the primitive, ourselves and the barbarian, ourselves and t

Fig. 2. Part of the concordance lines of his, our, ourselves

*Concordance* refers a tool which searches a text or a corpus for a selected word or phrase (a node) and presents every instance of the node in the center, with words that come before and after the node presented to the left and right. Concordance lines are the instances the tool presents in accordance with the nodes keyed in. Some corpus tools such as AntConc provides a function which allows users to sort and highlight words of different distance to the left or right of the node. Figure 2 shows what concordance lines look like.

### 3.4 The Teaching Experiment

In the teaching experiment, there were four lessons each week, each lasting for 40 min. The first week was used to demonstrate how corpus-aided analytical reading was carried out. In the following four weeks, the students studied four passages using the corpus-aided DDL method. In the last week, the questionnaire survey and interviews were carried out. The corpus-aided teaching method consisted of the following steps:

1. Prepare a corpus of the reading passages ready to be searched by AntConc.
2. Prepare the students with the tools used, reviewing the functions of AntConc, i.e. *wordlist*, *keywordlist*, *concordances*, and *collocates*.
3. Demonstrate to the students how to make keyword lists (the content words with significant keyness indicated in Fig. 1) and guide them to analyze and categorize the words, as shown in Table 2.

**Table 2.** Keywords identified and categorized from The Science of Custom

Noun						Adjective		Verb	
Type	Token	Type	Token	Type	Token	Type	Token	Type	Token
custom(s)	13	part	3	institutions	2	social	7	born	3
study	8	human	3	impossibilities	2	own	7	weighting	2
behavior	5	cultures	2	forms	2	individual	3	group	2
Anthropology (gist)	5	role	2	experience	2	necessary	3		
belief(s)	4	varieties	2	community	2	possible	3		
matter	4	sciences	2	civilization	2	traditional	2		
habits	3	propositions	2			interested	2		
tradition	3	laws	2						
Word types 22; Tokens 75						Word types 7; Tokens 27		Word types 3; Tokens 7	
Total word types 32; Total tokens 109									

4. Guide students to make prediction, and discuss possible connections, logic relationships of the items on the keywords list.
5. Have students’ attention focus on the most important keywords and use them as nodes to work out the concordance lines for analytical reading. Figure 3 shows part of the concordance lines of *custom*.<sup>1</sup> Figure 4 shows how guided reading on the words and phrases associating with the nodes is done.

```

2 eat gamut of custom that is found in various cultures, and hi
3 in which the customs of any peoples function in the lives of
5 igation, but custom, we have a way of thinking, is behavior a
6 Traditional custom, taken the world over, is a mass of detai
7 nt role that custom plays in experience and in belief, and th
8 inite set of customs and institutions and ways of thinking. E
10 rt played by custom in shaping the behavior of the individual
11 traditional custom, is as the proportion of the total vocabu
12 is birth the customs into which he is born shape his experien
    
```

**Fig. 3.** Part of the concordance lines of *custom*\*

<sup>1</sup> In AntConc, \* is a wildcard for zero or more characters. *Custom*\* therefore includes *custom* and *customs*.

1 Anthropology is the study of human beings as creatures  
 2 opology is the study of human beings as creatures of society. It fast  
 3 ishing mark of anthropology among the social sciences is that it incl  
 4 es for serious study other societies than our own. For its purposes a  
 5 zation. To the anthropologist, our customs and those of a New Guinea  
 6 he remains an anthropologist he is bound to avoid any weighting of o  
 7 one seriously studies social orders that have had the opportunity to  
 8 lligible. The study of custom can be profitable only after certain p  
 9 any scientific study requires that there be no preferential weighting  
 10 ields like the study of cacti or termites or the nature of nebulae, t  
 11 sary method of study is to group the relevant material and to take no  
 12 is only in the study of man himself that the major social sciences ha  
 13 substituted the study of one local variation that of Western civiliza  
 14 civilization. Anthropology was by definition impossible as long as t

Fig. 4. Part of the concordance lines of *study* and *anthropolog*\*

6. Encourage students to use corpus tools to read for key information, identify how the information is organized, and find out the deep meanings underneath the text, etc. by helping them understand the theme and its supporting sub-themes, meanings contained in keywords, and their possible logical connections.
7. Have students read the text and bear the analysis they have done with the corpus data in mind to confirm and deepen their understanding.
8. After Week one's demonstration, students were required to do the corpus-aided DDL reading for four weeks with necessary teacher guidance.

After the teaching experiment, a questionnaire survey was carried out to gather responses from the students. Part One of the survey attempted to find out to the students' opinion about the effect of the teaching method, and Part Two focused on their suggestion for the use of the method by having them answer a multiple-choice question (see Table 4). On the basis of the survey, follow-up interviews were conducted with some of the students.

### 3.5 The Reading-Writing Experiment

The survey and interviews after the teaching experiment suggests that the students expected to have the corpus processed input prior to reading, and that they would like to have teacher guidance for using the corpus materials. To further investigate the significance of teacher instruction in DDL, a reading-writing experiment was conducted. Three classes of English-major students from the same cohort (classes A, B, and C) were required to read a passage with different methods of instruction and to write a summary of it at the end of the experiment.

The reading passage used in the experiment was a 1674-word essay *Some Thoughts on Writing*<sup>2</sup>. The additional corpus materials included a keyword list and concordance lines of selected keywords of the passage. AntConc was used to compare the reading passage with the rest of the same textbook (79,960 words). The theme of the passage on writing can be inferred from some of most salient keywords such as *writing, write, work, how, published, rejected*, etc. The frequent mention of first-person and second-

<sup>2</sup> The article was written by Elizabeth Gilbert and was adapted in *A New English Course Book 4* unit 3 intended for second-year English majors.

person pronouns such as *I, my, you, your*, etc. is a feature of the dialogic style of the writing. In addition, concordance lines were extracted for three selected keywords, *write, you*, and *I as well as* their inflectional forms.

The methods of instruction involved two variables: form of input and teacher instruction. Class A read the passage with teacher instruction. Class B were given both the passage and corpus-aided materials, but no teacher instruction. Class C were provided with the passage as well as the corpus input with teacher instruction. The instruction provided to Class A was regular directions in a reading lesson and brief guidance about the topic and organization of the text, while in Class C it focused on the reading of the concordance lines, especially on guiding the students' attention to words and expressions indicating the topic and subtopics of the reading passage.

The experiment was carried out in language labs. It lasted for 40 min, within which the reading process and the writing task were completed. The summaries were written directly on computer and were collected at the end of the experiment. Among the 135 pieces of summary collected, 16 of them being too short or not in the proper form of summary were excluded. The remaining 119 valid copies of summary were used as the final data of the study. Table 3 provides description of the summaries collected.

**Table 3.** Description of the summaries collected

Group	Instruction		No. of text	Token	Type
	corpus-aided input	Teacher instruction			
Class A	–	+	42	4374	667
Class B	+	–	34	5368	885
Class C	+	+	43	5947	786

After the summaries were gathered they were carefully graded with the following criteria: (1) *Main idea coverage* which focuses on the number of main ideas included in the written summary; (2) *Integration* which examines the extent to which the information in the text is presented succinctly; and (3) *Language use* which focuses on correctness of grammar and vocabulary use. Four independent raters were invited to do the rating and an average score of each summary was obtained.

## 4 Results and Discussion

The results of the questionnaire survey are shown in Table 4. The results from Part One show that the students had a positive attitude towards the teaching method. 93.4% of them agreed that the corpus-aided method improved their understanding of the theme and main ideas. 75.8% reported that the method was more likely to attract their attention to the language forms. 83.3% stated that the method deepened their understanding of the meaning of the keywords and their contribution to the logic development of the texts.



**Table 4.** Results of the survey

<i>Part One</i>						
Item	A Strongly agree No./%	B Agree No./%	C Not sure No./ %	D Disagree No./%	E Strongly disagree No./%	Total number and %
1. The corpus aided teaching can help me improve my understanding of the theme and main ideas better	8/26.7	20/66.7	1/3.3	1/3.3	0/0	30/100%
2. The corpus aided teaching is more likely to attract my attention to the language forms	5/17.2	17/58.6	5/17.2	2/6.9	0/0	29/100
3. It helps me deepen my understanding of what the keywords mean and how they contribute to the logic development of the texts	4/13.3	21/70	5/16.7	0/0	0/0	30/100
<i>Part Two</i>						
What would you like to see more for better use of the method? (multiple choices)	A More introduction of the corpus tools and techniques					16/77 20.8%
	B Use it before analyzing the text in detail					18/77 23.4%
	C Use it after analyzing the text in detail					7/77 9.1%
	D Print the materials out for reading					11/77 14.3%
	E Teacher guided reading					25/77 32.5%

The data of Part Two indicate that the students had a higher expectation for teacher guided reading with corpus-aided input. It is the top choice they selected, taking up a third of all the choices they made. Compared with other aspects such as further explanation of the use of the tools, whether printed input is necessary and the time to use the corpus-aided materials, the students regarded teacher guidance to be more important.

The reason students positively commented on the corpus-aided DDL in reading is due to the fact that they were more engaged in reading. While they explored the theme of the text and its development, they were given bottom-up data centering on the keyness of a text. The corpus tools which instantaneously highlight the keywords, and the different neighbouring words of the nodes in concordance lines can help focus students’ attention on forms and observe how meanings are constructed. In the follow-up interviews, students reported that they were able to “quickly grasp the main theme and relationships of topic and sub-topics” while in traditional reading class they “tend to read word by word or line by line from the beginning to the end and still miss the core meaning”. They stated that “exploration of the texts is more effective because specific tasks are given”, and that “the rich data displayed and the requirements to work out the themes and the underneath meanings help to improve my critical reading ability”.

With regard to the reading-writing experiment, the treatment of both corpus-aided input and teacher instruction proved to be most effective in helping the students with summary writing. Table 5 shows the results of the experiment.

**Table 5.** Statistics of results of the reading-writing experiment

Descriptives						
Score						
	N	Mean	Std. Deviation	Std. Error	95% confidence interval for mean	
					Lower bound	Upper Bound
Class A	42	76.5476	6.65693	1.02719	74.4732	78.6221
Class B	34	79.5294	7.68996	1.31882	76.8463	82.2126
Class C	43	81.1047	7.51155	1.14550	78.7929	83.4164
Total	119	79.0462	7.47340	.68509	77.6896	80.4029
ANOVA						
Score						
	Sum of squares	df	Mean square	F	Sig.	
Between groups	452.341	2	226.171	4.274	.016	
Within groups	6138.154	116	52.915			
Total	6590.496	118				

The mean scores of summary writing of classes A, B, and C are 76.55, 79.53 and 81.1 respectively. Classes B and C with corpus-aided input performed better than Class A who received no corpus-aided input. Class C with both corpus-aided input and teacher guidance performed the best. Its mean score is significantly higher than that of Class A, suggesting that teacher-guided use of corpus-aided materials is effective in enhancing analytical reading.

Summary writing is an important output skill which is closely related to one’s reading quality. A good summary includes the main idea of the text, the most essential

supporting details and the use of paraphrasing skills with correct grammar and vocabulary. The skill to handle these elements in summary writing has much to do with the students' critical reading ability which lays a foundation for the writing. The experiment provides evidence of the positive influence of the mode of corpus-aided input with teacher guidance on analytical reading. It also indicates the importance of considering learner needs in teaching, namely, in-depth interaction with bottom-up data will be more effective when coupled with proper guidance and challenge from teacher-student interactions.

## 5 Concluding Remarks

The present study constructed a 6-week reading teaching experiment with Chinese advanced EFL learners using the corpus-aided DDL method. The effect of the method was investigated through a questionnaire survey, interviews and a follow-up reading-writing experiment. Different from previous studies which used much larger corpus when applying the DDL approach, the study concerns single texts. It shows how corpus tools can be used to assist analytical reading and language learning. The corpus-aided bottom-up method helps students to focus on the most relevant information of a text. When students observe keywords, concordances, clusters centering on the nodes, etc., they are exploring how the theme and its sub-topics or supporting details developed. When reading keywordlists, for example, they make comparison and categorization of the contents. When observing concordance lines of selected nodes, they make inference and evaluation of the ideas expressed.

The integration of technology with teaching brings about chances for classroom teaching innovation. The present study suggests that the effect of such innovations can be mediated by the teacher instruction provided. In addition, it provides implications for further research applying the corpus-aided DDL method to reading teaching of more texts in one particular discipline or subject area.

**Acknowledgements.** This paper is the result of the following funds: Guangdong Higher Education Teaching Reform Project 2016 (No. 236); Guangdong "13th Five Year" Plan Co-funded Project of Philosophy & Social Science (GD16XWW25). The corresponding author of this paper is Xiaobin Liu.

## References

- Anthony, L.: AntConc (Version 3.2.0) [Computer Software]. Tokyo, Japan: Waseda University (2014). <http://www.laurenceanthony.net/software>
- Bernardini, S.: Corpora in the classroom: an overview and some reflections on future developments. In: Sinclair, J.M. (ed.) *How to Use Corpora in Language Teaching*, pp. 15–38. John Benjamins, Amsterdam (2004)
- Boekaerts, M., Corno, L.: Self-regulation in the classroom: a perspective on assessment and intervention. *Appl. Psychol.* **54**(2), 199–231 (2005)
- Boulton, A., Cobb, T.: Corpus use in language learning: a meta-analysis. *Lang. Learn.* **67**(2), 348–393 (2017)

- Chall, J.: *Stages of Reading Development*. McGraw-Hill, New York (1983)
- Chen, M., Flowerdew, J.: Introducing data-driven learning to PhD students for research writing purposes: a territory-wide project in Hong Kong. *Engl. Specif. Purp.* **50**, 97–112 (2018)
- Cobb, T., Boulton, A.: Classroom applications of corpus analysis. In: Biber, D., Reppen, R. (eds.) *The Cambridge Handbook of English Corpus Linguistics*, pp. 478–497. Cambridge University Press, Cambridge (2015)
- Gabrielatos, C.: Corpora and language teaching: just a fling or wedding bells? *Tesl-Ej* **8**(4), n4 (2005)
- Gui, S., et al.: Corpus linguistics and China's foreign language teaching. *Mod. Foreign Lang.* **4**, 419–426 (2010)
- Hadley, G., Charles, M.: Enhancing extensive reading with data-driven learning. *Lang. Learn. Technol.* **21**(3), 131–152 (2017)
- He, A.: A review on the development of corpus processing for teaching. *Foreign Lang. China* **4**, 47–52 (2010)
- Jamshidifarsani, H., Garbaya, S., Lim, T., Blazevic, P., Ritchie, J.M.: Technology-based reading intervention programs for elementary grades: an analytical review. *Comput. Educ.* **128**, 427–451 (2019)
- Johns, T.: Should you be persuaded: Two samples of data-driven learning materials. In: Johns, T., King, P. (eds.) *Classroom Concordancing: English Language Research Journal* 4, pp. 1–16. University of Birmingham, Centre for English Language Studies, Birmingham, UK (1991)
- Kamil, M.L., Chou, H.K.: 13 comprehension and computer technology. In: Israel, S.E., Duffy, G. G. (eds.) *Handbook of Research on Reading Comprehension*, p. 289. Routledge, New York (2009)
- Larsen-Walker, M.: Can Data Driven Learning address L2 writers' habitual errors with English linking adverbials? *System* **69**, 26–37 (2017)
- McEnery, T., Xiao, R., Tono, Y.: *Corpus-Based Language Studies: An Advanced Resource Book*. Routledge, New York (2006)
- Scott, M., Tribble, C.: *Textual Patterns: Key Words and Corpus Analysis in Language Education*. John Benjamins Publishing Company, Amsterdam (2006)
- Rutherford, W.E.: *Second Language Grammar: Learning and Teaching*. Routledge, New York (2014)
- Zimmerman, B.J., Tsikalas, K.E.: Can computer-based learning environments (CBLEs) be used as self-regulatory tools to enhance learning? *Educ. Psychol.* **40**(4), 267–271 (2005)
- 高等学校外语专业教学指导委员会英语组. 高等学校英语专业英语教学大纲 [Z]. 北京: 外语教学与研究出版社; 上海: 上海外语教育出版社, 2000. Higher Education Advisor Committee of Foreign Language Teaching (the English Group). (2000). *Syllabus of English Teaching for English Majors of Higher Education*. Shanghai, China: Shanghai Foreign Language Education Press