

Vocabulary-supported Writing: Concordance-based Writing

Hyonjung Kim (Seoul Women's University)

Dongkwang Shin (Korea Institute of Curriculum & Evaluation)

I. Introduction

At the initial stage of the corpus-based technology development, corpora were mainly used for developing syllabi, instructional materials and reference works typically by analyzing the most frequent items found in the corpora. However, as corpus-based skills have surprisingly evolved in eighties and nineties, the focus has moved onto an online-based model. In this model, raw concordances derived from corpora are directly used to exemplify real-life usage of a target item. This approach has been promoted as “data-driven learning” by Tim Johns (1986, 1988, 1991).

Data-driven learning focuses on discovery activities by which learners inductively derive and deductively apply their hypotheses by classifying data from corpora. In most EFL classroom situations, the amount of real English input available for learners is usually very limited. Data-driven learning using corpora can provide with a concentrated source of English input. This procedure is based on the effectiveness of inductive learning from multiple examples (Ellis, 1996; Skehan, 1998).

The concordancer derives all the contexts available for a target word or phrase from a text or corpus, which cannot ensure all the concordance samples are helpful for learners to guess the meaning of a target word. Nevertheless, one or more of the given contexts is likely to provide appropriate learning conditions needed by a particular learner to understand the meaning of a target word and real uses of the word (Cobb, Greaves, & Horst, 2001).

In the present study, the basic assumption is that those concordancers can guide learners to produce $i+1$ output beyond their current grammatical competence, thus they can more concentrate on content. We will examine how effectively “data-driven writing” can contribute to the quality of writing.

There are two research questions in the present study; 1) To what degree can data-driven writing contribute to reducing grammatical errors in writing, and 2) How effective is using corpus-based skills for administering learners' writing drafts?

II. The study

For the pilot test three first-year college students participated in the study; two from advanced and one from intermediate level. They were all non-English majors. The students were required to write a short essay on a given topic about a page long. Then they were asked to revise their essay and hand it in. After their second draft, the students were introduced to Yahoo Mini (http://kr.dic.yahoo.com/search/mini/mini_yahoo_download.html) and were trained how to use the Concordance Writer (<http://www.er.uqam.ca/nobel/r21270/concordwriter/>). The final draft is the result of incorporating both computer programs in their writings.

All errors the subjects made are given a unique tag. For example, [PRE] represents a preposition error. By analyzing the error-tagged drafts through the RANGE program (<http://www.vuw.ac.nz/lals/staff/paul-nation/nation.aspx>), we can count the number of all errors or a particular error type at once. Thus, we can check what kind of error is the most frequent in a particular draft and what kind error has increased or decreased in drafts by a particular subject for the experiment period. In addition, if we use the concordance AntConc (<http://www.antlab.sci.waseda.ac.jp/software/antconc3.2.1w.exe>) for analyzing the drafts, we can look for all contexts related to a particular error type.

III. Results and Discussion

The analysis using the RANGE program showed that the students' errors decreased over their rewriting processes. However, we cannot be conclusive whether using concordance samples was the main reason for the reduction in their errors since the decrease in the students' error could have been resulted from their rewriting process itself, not due to the influence of using Concord Writer or Yahoo Mini.

Further interview with the students revealed interesting insights in the study. It was found that the intermediate student could not apply Concord Writer in her writing due to her limited English proficiency. First of all, she could not identify her own errors and second, even if she did, she was unable to understand the sample given in Concord Writer to correct her own writing. Thus, she mostly used Yahoo Mini, especially Korean-English dictionary to look up words that she wanted to express in English. The advanced, however, found both the Concord Writer and Yahoo Mini useful. This was because they were proficient enough to recognize, check and thus correct their errors using those two programs. For further research it is recommended to use a larger corpus for examining whether the corpus size affect the writing skill in the advanced student.

References

- Cobb, T., Greaves, C., & Horst, M. (2001). Can the rate of lexical acquisition from reading be increased? An experiment in reading French with a suite of on-line resources. In P. Raymond & C. Cornaire (Eds.), *Regards sur la didactique des langues secondes* (pp. 133-153). Montréal: Éditions logique.
- Ellis, N. C. (1996). Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition*, 18, 91-126.
- Johns, T. (1986). Micro-concord: A language learner's research tool. *System*, 14(2), 151-162.
- Johns, T. (1988). Whence and whither classroom concordancing? In T. Bongaerts, P. de Haan, S. Lobbe & H. Wekker (Eds.), *Computer applications in language learning* (pp. 9-27). London: Foris.
- Johns, T. (1991). Should you be persuaded - two examples of data-driven learning. In T. Johns & P. King (Eds.), *Classroom concordancing: English Language Research Journal*, 4 (pp. 1-13). University of Birmingham: Centre for English Language Studies.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Appendices

In the following tables, DTE signals Determiner Error, NE for Number (Subject-Verb) Agreement Error, PE for Performance Error, PRE for Preposition Error, PSE for Possessive Case Error, VOE for Verb Omission Error, VAE for Verb Appropriateness Error and CE for Citation Error.

1. Sample A (High)

TYPE	RANGE	FREQ	F1	F2	F3
DTE	3	11	4	5	2
NE	1	2	2	0	0
PE	3	12	6	3	3
PRE	3	3	1	1	1
PSE	3	3	1	1	1
VOE	3	5	2	1	2
Sum			16	11	9

2. Sample B (Intermediate)

TYPE	RANGE	FREQ	F1	F2	F3

DTE	3	15	6	4	5
VAE	2	4	2	2	0
Sum			8	6	5

3. Sample C (High)

TYPE	RANGE	FREQ	F1	F2	F3
CE	2	2	1	1	1
DTE	3	4	2	1	1
VAE	3	3	1	1	1
Sum			4	3	3

Hyonjung Kim

E-mail: kayli@swu.ac.kr

Phone: 02-970-7875

Dongkwang Shin

E-mail: sdhera@kice.re.kr

Phone: 02-3704-3816