

# The Application of Corpus in English Composition Teaching

Jing Chen<sup>+</sup>, Lingyan Xiong

Faculty of Foreign Studies, Jiangxi University of Science and Technology,  
Ganzhou, 341000, China

**Abstract.** As we all know, computer can complete what traditional research cannot do. Computer corpus skill applying to the linguistic research and teaching process increasingly, taking Writer's Workbench (WWB) for example, the author try to analysis corpus's applying to English composition and the function of the building of the small size specialized corpus hoping to bring some changes for traditional model of college English composition teaching. Facts show us the corpus software, WWB, is very powerful, having wide usage and a teaching and research function.

**Keywords:** Corpus, Second Language Acquisition, College English Reform, Composition Model

## 1. Introduction

Corpus is the study of language based on examples of "real life" language use stored in corpora (or corpuses)—computerized databases created for linguistic research. With the hyper growth of computer skill and corpus research and application in processing and analysis, recently, second language acquisition show a diverse depolarization modern developing trends [1]. The corpus based on computer technique can use language statistical analysis technique to have an objective and scientific measure on learners' language use. Corpus, large capacity for large sample material true, quantitative statistics analysis and case qualitative text analysis provides may, and can long-term preservation copy retrieval, for verification research to provide the possibility [2].

Making use of computer technique, corpus can easily to have a series of operation of analysis, input and mark. That has greatly improved the corpora collection scope and the processing speed. The application of retrieval technology provides essential condition for corpus. Using corpus retrieval tools, we can study context concurrence, statistical frequency and so on. Along with the development of the corpus application, the corpus value in system retrieval data and inspection linguistics assumption has been more and more recognized by people [3].

The third document published by Department of academic education points out the teaching mode reform is one of the purposes of promoting students' individualized learning method to the formation and development of the student's ability of autonomous learning. New teaching mode should enable the students to choose their own suitable required materials and methods of learning, obtaining learning strategies of guidance, and gradually raise their autonomous learning ability. To college English, the application of corpus in English writing teaching is an important revolution for the traditional college English writing teaching mode, adopting diversified forms, experimental new teaching auxiliary tools, promoting the improvement of students' autonomous learning ability, practicing teaching of interactivity participation. It has changed the college students' learning environment and the process of writing, promote the improvement of students' autonomous learning ability. For many English teachers, the correction of composition and marking is time-consuming and hard work, so I conduct a study on the developed text analysis program Writer's Workbench (hereinafter referred to as WWB) developed by AT& T Bell LABS, so as to use corpus as an auxiliary for writing teaching helping improve students' writing ability. WWB software installation can hind as a plugin of

---

<sup>+</sup> Corresponding author. Tel.: +013607974281.  
E-mail address: cjxf99521@163.com.

Microsoft Word, used with Word software at the same time, analyzing directly and providing immediate feedback to the Word doc writing text. WWB can serve for diagnosis type teaching, judging composition quickly in a short time, searching and analyzing the problem of vocabulary, sentence structure, format, breaks, cohesion ways, giving evaluation to the quality of composition [4]. It is also good for the students to use for they can pay more attention to the linguistic characteristics of text under the guidance of teachers.

By the way, let me introduce the function of the software WWB briefly:

This software is very powerful, having wide usage and a teaching and research function. The structure of the software includes 25 analysis programs which were divided into 6 categories: content, verbs, clarity, words, punctuation. Each category has different colors. Users click on any distinguish button, the system will automatically report the results of analysis in the forms of new files.

## 2. The construction and utilization of corpus

Before building small corpus, several basic problems should be clear: what kind of corpora should be collected? What kind of problem should be expected to analyze? What is the aim? This study is about a small corpus based on college English writing to find out the shortcomings and errors in writing by analyzing the mistakes on words, phrases and sentences appeared in composition. And in order to improve students' English writing ability, we would like to provide available corpus resources to realize the diversification of subject and phrase.

## 3. Collecting corpus resources

Usually I collected college students' compositions as the corpus resources. After writing compositions, students are required to submit up them in forms of electronic paper. The untreated corpus materials would be put into separate files, depositing after named.

## 4. Analysis examples——language characteristics

Language characteristics are the focus of writing teaching. WWB language feature analysis can analyze prose style, including sentence length and style, the use of verbs, the beginning of an article, the legibility indices of an article, and so on. After analyzing a student's article entitled, "education is a life-long progress" using WWB, we reach the following results (Figure 1).

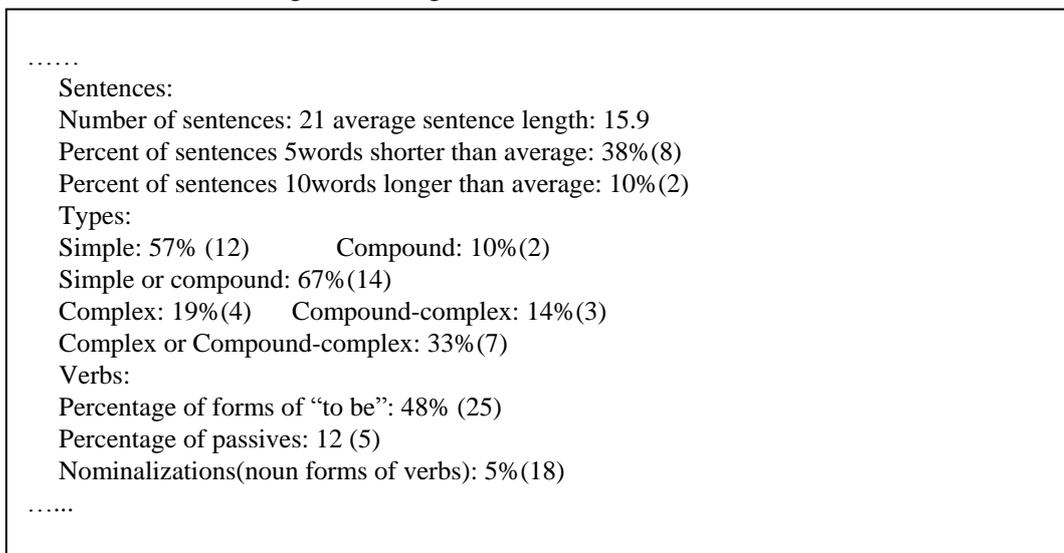


Fig. 1. Example of language characteristics.

Analysis results show that, the average length of the sentence in the composition is 15.9 words, lower than the set standard (18); the percentage of the sentences whose length is less than the average sentence length 5 words occupies 38%, which is higher than the standard(30%); the percentage of those whose length is more than the average sentence length 10 words is 10%, which is lower than the standard(15%). Thus, it can be seen, the sentence length overall in this composition is short.

According to sentence types, simple sentences and parallel sentences occupy 67% of all the sentences, substantially higher than the standard (50%); the compound sentences and parallel compound sentences occupy 33%, greatly less than 50%. Thus, it can be seen, the composition tends to use simple sentences.

In the verb forms, the composition used the form "to be" 25 times, having very high frequency (occupy 48% of all the verbs of the composition); the passive sentences occupy 12%, close to 10% of the standard; the verb nominalization form of 5%, accord with less than 10% of the standard.

By above we reach the following conclusion that the average sentence length of this article is less than 18 words and long sentences are scarce, therefore we need to merge some sentences; the form "to be" is used too frequently and we should reduce the using of verbs nominalization such as "education".

## 5. Conclusion

Promoting effect of corpus in EFL teaching is obvious. Above case is only the iceberg of corpus applying to teaching research. Introducing English corpus into classroom teaching can narrow the distance between the language taught in class and the native language, cultivate students' language awareness and autonomous learning ability. And the corpus research on the interact words among teachers and students can carry out effective monitoring on the classroom discourse output, providing the basis for improving classroom teaching. The WWB application is varied. Other software, for example, Wordsmith, Concordance, Paraconc etc are widely used in second language acquisition research. In teaching, corpus can help us set scientifically teaching contents and procedure, extending classroom teaching space such as used in college English writing teaching, can change the past teaching pattern from forms and thinking, reducing teachers' burden of composition correction, and meanwhile make teaching content more rich. Students can also decide independently their own study content and schedule, evaluate their study effect, manage their learning process, be responsible to their own learning behavior, make writing class become really student-centered. Some experts were highly valued corpus application to foreign language class teaching, however, corpus application in the foreign language teaching is still pending reclaiming fields. We need to have a batch of teaching personnel to engage in the reform of the traditional teaching models, exploring boldly the teaching implementation of corpus, making foreign language teaching program more scientific and reasonable.

## 6. Acknowledgment

The study on which this paper was based was supported by the Science Foundation of Jiangxi University of Science and Technology (No. jxxj11110) and the Science Foundation of Ganzhou Social Science Federation (2009). Authors gratefully thank the anonymous reviewers for their valuable comments on this manuscript.

I would like to acknowledge the contributions of the following groups and individuals to the development of my essay:

Ms. Librarian at Jiangxi University of Science and Technology

Ms. Librarian directed me to a wide range of resources on the web and in the library stacks. She answered all of my questions patiently as well as asked me questions that helped me to narrow my search. Further, she helped me figure out correct documentation for sources.

My class peer research group: LingYan Xiong and Xiaoping Zou

Part of the students' compositions I used as my corpus materials are provided by them.

## 7. References

- [1] Krashen, S.D. *Second Language Acquisition and Second Language Learning*. Pergamon Press. 1981. 87.
- [2] Ministry of Education Office. College English curriculum requirements of the notice 2007.
- [3] Yang Huizhong. *Corpus linguistics introduction*. Shanghai: Shanghai foreign language education press, 2002. 16
- [4] Wang Lifei computer aided second language research methods and application. Foreign language teaching and research press, 2007.64.