The Use of Modal Verb in English Writing: Corpus Views

Deqing Wang

Depart. of Culture, Shandong Art and Culture School, Jinan Shandong, China

deqing1123@163.com

Abstract. This study analyzes the use of modal verb based on the corpus of Chinese English learners. The purpose of the study is to investigate the situation of modal verb used by English majors in China in their writing. Through the application of advanced software, this study reaches the results that the frequency of the English majors using modal verb is phased and associated with students' language competence, but differs greatly. However, the students who have low level of English tend to use the modal verb.

Introduction

Modal verbs have been a heated topic in writing circle during the past 30 years. By now, it is very common to use a corpus to investigate the application of modal verb. It seems that students need to take notice of modal verb. Learners try to do the research, but modal verb got little attention. This thesis seeks to remedy this omission, applying an approach to modal verbs in English by comparing the native corpus with the learner corpus. The study of modal verbs used by the corpus is more objective and comprehensive. This study aims to find out the using tendency of modal verb English majors in writing by comparing the native English corpus with the Chinese learner corpus.

The motivation of this paper is like this: Firstly, modal verb is helpful to encourage the author to write this thesis. The modal verbs are very important to the writing of English. Students can use the modal verb to know the sequences of Language. But we know about our world so little that we need to use some words and to talk about what may be true and may be not, what ought to be and ought not to be, what certain individuals are capable of and what is impossible for them, what obligations we have to do or refrain from doing.

Secondly, there is a certain degree of practical significance, our English teaching and learning. Hope that the results of this study will reveal the modal sequence common problems in ESL/EFL writing environment. At the same time, also hope that, by the high score group and low score group are established using modal series automatic scoring system useful.

This study can help us to know the use of modal verbs in English writing by Chinese English learners. Also, it is helpful to improve English writing skills.

Literature Review

Corpus Linguistics. Corpus linguistics can best be defined as a linguistic methodology which is founded on the use of electronic collections of naturally occurring texts, viz. corpora. This is a new research subject of enterprise, a new method of philosophy, an open sesame in a new way of thinking about language (Davis, 2001). Computer software tool combined with the number of language data as evidence reveals and will continue to reveal previously unknown language phenomenon and the diversity of power is impressive.

For Sinclair (1996: 76) the heuristic power of corpus methods is no longer in doubt. Applied linguistics in recent works is based on the specific suggestions on the results of theoretical research, practice and language in the real world, but also the defects. Another method is to directly study the language structure and use them in the real world domain model, so that the application of the

proposal is based on a solid empirical foundation. Based on the analysis of corpora is especially for this purpose.

Previous Study of Modals. A lot of studies have been studied from many point on the modals : speech acts (Boyd & Thome, 1969: 57-74), systemic grammar (Bybee et al., 1995), feature analysis (Marino, 1973: 309-323) and of stratification grammar (Granger, 2002). These many varied approaches have failed because of their reliance on the analyst's intuition (Coates, 1983).

Many researches have started to figure out something about modal auxiliaries (e.g. Coates 1983; Kennedy, 2002). These studies have tended to focus on the following questions:

(1)What is the relative frequency of the various modal auxiliaries in naturally-occurring English, and in learner English?

(2)What is the outcome of meaning between the modals in native English and in learner English?

Major corpus-based empirical studies fall mainly into three groups in line with different types of corpora in use:

(1) The investigation of native corpus, especially British or American corpus (e.g. Coates, 1983; Hunston, 2004; Kennedy, 2008);

(2) Comparison between native corpora and non-native learner corpora (e.g. Neff et. al., 2003; Römer, 2004);

Studies on the Use of Modal Verbs by Chinese Native Speakers. The first corpus-based study of modal verbs is Ehrman (1966) he studied the American modal system with the Brown Corpus. Coates (1983), and Perkins (1983) did research on British English corpora basing on samples of written and spoken English, Quite a lot of research analyze the use of English central modals in legal or academic texts (Neff et al., 2003; Hunston, 2004) to show that central modal verb have their own ways of distribution in writing.

Studies on the Use of Modal Verbs by EFL/ESL Learners. Aijmer (2002) built corpus to compare the use of modal words in native English writing and EFL writing of advanced university students. Neff et al. (2003) came out results that the EFL writers overuse modal verbs when it comes to in comparison with American writers, and the can is overused by all non-native writers.

Römer (2004) makes a comparative analysis of modal auxiliaries, including their distributions, meanings and contexts in the spoken part of British English Corpus—BNC data. In China, Liu Hua (2004: 121-125) makes a study on modal verbs based on native corpus—Freiburg Lancaster Oslo/Bergen corpus (FLOB) and the sub-corpus of Chinese Learner English Corpus (CLEC). Through his research he found out dare, should, can, will, have to, be to, need, must and may are overused by learners; while shall, might, could, would and ought are underused. Modal verbs were studied between the native corpus - the written part in BNC and the learner corpus - ST6 part in CLEC. The result is that English learners tend to use the modal verbs. The above phenomenon is called topic-dependency (Liang Maocheng, 2008: 51). It results that the more relied of the writing, the higher use of the same words. It results in the limited selection of English vocabulary.

These corpus-based studies concerning the range, frequency distribution of modal expressions lead to a problem in their studies. That is the corpus they chosen has big difference with the learner corpus. The study of word frequency has strong topic-dependency, so it deserves further research.

Methodology

The basic approach used in this paper is a corpus-based qualitative and quantitative research. In order to find out the modal sequences used in writing frequency of English Majors in China, this study uses the LOCNESS corpus, and Contrastive Interlanguage Analysis is adopted in this paper.

Research Questions. This paper aims to investigate the using tendency of written form of the modal sequences of English majors in their writing. To be more specific, four major questions are addressed:

1) How do English majors differ from native speakers in the use of modal verbs in universities?

2) How often English major students use modal verb? Such as would or should

Research Methods.

Methodology of Corpus-Based Approach. Corpus-based approach has become a mainstream in language study. Biber, Conrad, and Reppen (2008) summarize the essential characteristics of corpus-based analysis: it is empirical, making a search on the use of modal verb in books and English writing; it makes a large study on the collection words on books. Which is known as a "corpus" .It is used as the basis for analysis; it makes full use of extensive use of computers for analysis, using both automatic and interactive techniques; it depends on both quantitative and qualitative techniques. Of course its result is very reliable and supportive. It gives us an open view of the situation of using modal verbs.

As we can see from the obvious four features of analysis that is based on corpus, we easily know there is a hope to use computers to analyze the result and the patterns of language use, allowing the storage and analysis of a larger database of natural language.

It is found that the outcome cannot prove the frequency is closely related with the students' competence and it has its own shortcoming and we cannot avoid it. The limitation is that the pattern of samples for examination; what's more, even the example are representative, specific findings by qualitative approach "Whether they are true to be discovered by using the corpus is a problem." These two methods for corpus linguistics are closely related and complementary, qualitative analysis is possible to produce precision However, quantitative analysis is reliable and generalized.

The author has the intention to make analysis of the essays written by university students at present, finding out the difference of using the modal verb between learner corpus and native corpus, and then draws some instructive conclusions to speed up the FL/SL teaching and learning.

Contrastive Interlanguage Analysis. Comparative analysis of different from the traditional sense of the (CA), to compare the different language, this method compare / contrast is a native language is not English and native speakers do in a similar situation. This new method is called "a new type of calcium; Granger (1997) refers to the Contrastive Interlanguage Analysis for the Central Intelligence agency.

Through the corpus-based analysis and contrastive interlanguage analysis, this research seeks to investigate the general use of modal sequences by Chinese English majors and the difference of these modal sequences used between Chinese English majors and native college students. Moreover, we should be aware that qualitative analysis is also needed in the quantitative studies, as Biber et al (2008) point out "that corpus must go beyond simple counts of linguistic features, it is essential to include qualitative, functional interpretations of quantitative patterns". Therefore, besides showing various data we draw from corpora, we devote a large part of the discussion section in this paper to a systematic and multi-level qualitative analysis.

Research Design. This section describes the design of the study. It begins with the introduction of corpora used in this study and goes to the data selection. Then the software adopted in this study and grouping of corpus data are introduced. At last, the main steps of analysis are introduced.

The present study focuses on the analysis of the language data from the learners' corpus - CLEC and the native corpus - BNC and LOCNESS. They are reliable and we can rely on it to select the data.

Results and Discussion

This part mainly tells about the results concluded from both quantitative and qualitative analysis of the overused and underused modal verbs. It consists of four parts. Section one studies the use of modal verbs by comparing learner corpus and native corpus in their writing. Section two investigates the using situation of modal verbs. In section three, through analyzing the semantic type of typical modal sequences, further discussions is also carried out to find whether there is some relationship between the English majors' language proficiency and the use of these modal sequences. In the last section, the author makes a short conclusion.

The Overall Frequency of Modal Verbs. At first, in order to arrive at comparable frequencies across corpora of different size, normalized data are provided in Fig. 1.
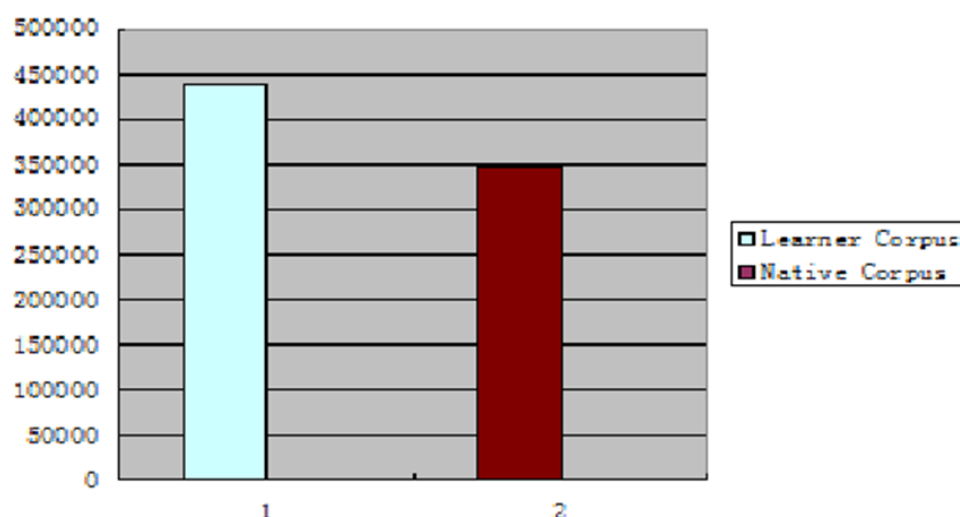
Figure 1. The Total Numbers of Overall Tokens in Learner Corpus and Native Corpus
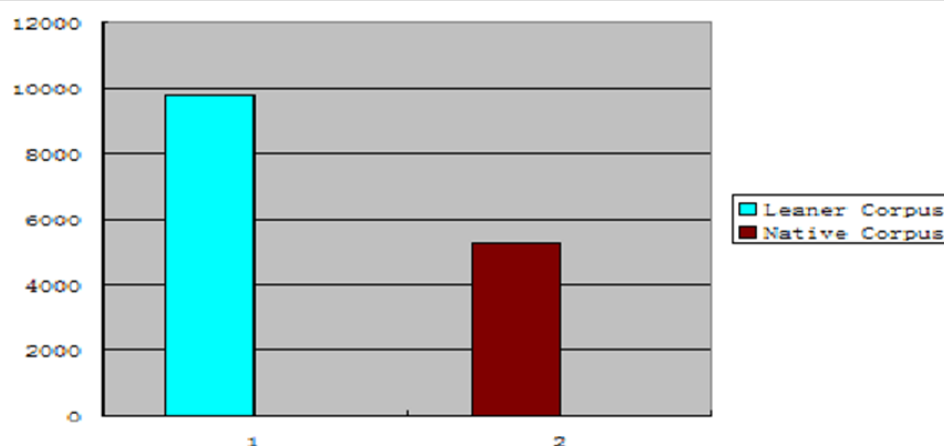

Figure 2. The Total Numbers of Overall Modal Verbs in Learner Corpus and Native Corpus

Fig. 2 demonstrates that the Chinese English majors like to overuse modal verbs. In other words, modal verbs in our English majors' writing are overused, and it is the same as Lu Xiaojuan (2007), the research of Liu Hua (2004), and Ma Gang. We also find that the results of Swedish English learners studied by Aijmer (2002) look like it. There are two reasons to clarify the problem. The first one is that college students are willing to use modal verb to express meaning. The second reason is that the modal verb is easy to adopt in English writing.

Table 1 The frequency of modal verbs in learner corpus and English major students corpus

Keywords	NNS		NS		Keyness	P Value
	frequency	‰	frequency	‰		
CAN	3304	7.50	1091	3.31	575.49	.000
WILL	2307	5.24	962	2.92	236.75	.000
SHOULD	1429	3.24	655	1.99	109.01	.000
MUST	396	0.90	291	0.88	−0.04	.000
COULD	440	1.00	536	1.63	−58.13	.000
WOULD	761	1.72	1193	3.62	−264.95	.000

The Overall Frequency of Modal Verbs. Table 1 apparently demonstrates that modal verbs in the native corpus are very different from these in the learner corpus native corpus by amount. It concludes that "can" are used most, "will" is the second largest number, and "should" is next to "will". These three modal verbs are overused, and their numbers are to7040 times and occupy 72.8% of the total modal verbs used in the selected learner corpus.

The Underuse of Modal Verbs. As we can see, "must", "could" and "should" are used little. Apart from those six modal verbs which are underused or overused mentioned above, there is little difference of frequency of "shall", "need", "ought to", "may", "might", "dare", "used to", which are partly due to the low frequency in both corpora. These seven modal verbs haven't got special features, so there is no further study in the following passages in this thesis.

The Comparison between English Majors and Non-English Majors. Liang Maocheng (2008) carried out a study on the use of modal verbs between non-English majors and native university students in their writing. He found the result that the use of other modal verbs in learner corpus and native corpus look like Liang Maocheng's finding showed in Table 2.

Table 2 The frequency of modal verbs in learner corpus and native corpus

Keywords	NNS		NS		Keyness	P Value
	frequency	‰	frequency	‰		
CAN	5298	13.8	1223	3.7	2179.45	.000
WILL	2926	7.6	916	2.8	825.45	.000
SHOULD	2242	5.8	653	2.0	699.24	.000
MUST	1596	4.1	282	0.9	827.03	.000
COULD	455	1.2	560	1.7	−32.86	.000
WOULD	486	1.3	1220	3.7	−449.35	.000

(Notes: the data is from Liang Maocheng's study. Based on the research of modal verb corpus by the Chinese English learner English)

Making a comparison between Table 1 and Table 2, we can see that "can" still get the highest number while "would" gets the lowest number. The frequency of "must" brought out big gap in the two tables. The frequency of "must" in non-English majors writing in learner corpus was much higher than that of the native corpus. The Number of "must" is 827 with the use of "must" by the English majors was the same to the native speakers. The number of "must" was only 0.04. That is no difference between English majors and native speakers about the frequency of "must". The word "much" in the Table 2 is higher than in Table 1. We had a conclusion that the overusing "can", "will", "should" by non-English majors English majors was more obvious than English majors .While the under use of "would" and "could" between English majors and non-English majors are alike.

Conclusion

This study pays more attention to the use of modal verb used by English majors on the basis of corpus. British National Corpus (BNC) and the sub-corpus of Chinese Learner English Corpus (CLEC) are selected to make a difference. We prove that the use of modal verb is related to the English writing competence. Through analyzing the date of modal verb used by English majors, teachers can provide some typical language phenomenon, from the main body of the students in their mastery of collocations and phrases. This method is to stimulate students' enthusiasm for learning in English learning and encourage them to study hard. A large corpus of LOCNESS like BNC, and can use the language through the careful design of the procedures of the real evidence.

English writing teaching has always been a perplexing problem, language teachers has been gradually formed in different education system bottlenecks in the training and improve their writing ability in English writing, a corpus is superior to the traditional teaching methods, which in turn provides an effective method to solve the problem of low efficiency and improve the teaching of English writing.

References

[1] Aijmer, K., Modality in advanced Swedish learners' written interlanguage, in: Granger et. al. (Eds.), Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching, John Benjamins Publishing Company, Amsterdam/Philadelphia, 2002, pp. 55-76.

[2] Biber, Conard and Reppen, Corpus Linguistics, Foreign Language Teaching and Research Press, Beijing, 2008.

[3] Boyd, J. & Thome, J. P., The semantics of modal verbs, J. Journal of Linguistics. 5 (1969) 57-74.

[4] Bybee, J. & S. Fleischman, Modality in grammar and discourse: An introductory essay, in: J. Bybee & S. Fleischman (Eds.), Modality in Grammar and Discourse, John Benjamins, Amsterdam, 1995, pp. 1-14.

[5] Coates, J., The Semantics of Modal Auxiliaries, Routledge, London, 1983.

[6] Davis, L., Doing culture: cross-culture aspects of second language writing, Shanghai Foreign Language Educational Press, Shanghai, 2001

[7] Ehrman, M., The Meanings of the Modals in Present Day American English, Mouton, The Hague, 1966.

[8] Granger, S., Automated retrieval of passives from native and learner corpora: Precision and recall, J. Journal of English Linguistics. 25(1997) 365-374.

[9] Granger, S., A Bird's eye view of learner corpus research, in: Granger et. al. (Eds.), Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching, John Benjamins Publishing Company, Amsterdam/Philadelphia, 2002, pp. 3-33.

[10] Kennedy, G., Variation in the distribution of modal verbs in the British National Corpus, in: R. Reppen, S. M. Fitzmaurice and D. Biber (Eds.), Using Corpora to Explore Linguistics Variation, Benjamins, Amsterdam, 2002, pp. 73-90.

[11] Kennedy, G., An Introduction to Corpus Linguistics, Longman, London, 2008.

[12] Hunston, S., "We can broke the forest", Approaches to modal auxiliaries in learner corpora, paper presented at TaLC6, Granada, July 2004.

[13] Liang Maocheng, A Corpus-based Study of Modal Sequences in Chinese Tertiary EFL Learners' Writing Production, J. Foreign Language Teaching and Research. 1 (2008) 51-58.

[14] Liu Hua, The Use of Modal Verbs by Senior English Majors in China, J. Journal of Ningbo University (Educational Science Edition). 5 (2004) 64-75.

[15] Ma Gang & Lu Xiaojuan, A Study on Modal Verbs based on CELC, Technology Enhanced Foreign Language Education. 3 (2007) 48-56.

[16] Neff, J., Dafouz, E., Herrera, H., Martinez, F & Rica, J.P., Contrasting learner corpora: the use of modal and reporting verbs in the expression of writer stance, J. Language and Computers. 48 (2003) 211-230.

[17] Palmer, F.R., Modality and the English modals, Longman, London, 1990.

[18] Perkins, M., Modal expressions in English, Frances Printer publishers, London, 1983.

[19] R ömer, U., A corpus-driven approach to modal auxiliaries and their didactics, in: J. M. Sinclair (Eds.), How to Use Corpora in Language Teaching, John Benjamins Publishing Company, Amsterdam/Philadelphia, 2004, pp. 185-199.

[20] Sinclair, J., The search for units of meaning, J. Textus. 1 (1996) 75-106.

[21] Tagliamonte, S., Have to, gotta, must: grammaticalisation, variation and specialization in English deontic modality, in: H. Lindquist and C. Mair (eds.), Corpus Approaches to Grammaticalization in English, Benjamins, Amsterdam, 2004, pp. 34-55.