

A Comparative Corpus-based Study of Semantic Prosody

Changhu Zhang

Dept. of Foreign Languages, Dezhou University, Dezhou, Shandong, China

E-mail: jackzch@126.com

Abstract— The present study uses the main corpus CLEC and the reference corpus BROWN to explore the semantic prosody of COMMIT in Chinese EFL. The result indicates that Chinese EFL learners exhibit similar semantic prosody as compared with those of native speakers. Nevertheless, they still use lots of interlanguage collocations and unusual collocations, which ruin the due semantic prosodic harmony, thus making their English quite unnatural and less idiomatic. Finally, implications on integrating semantic prosody into ESL/EFL vocabulary pedagogy are discussed.

Index Terms— semantic prosody, collocation, corpus, comparative

I. INTRODUCTION

In the past few years, a novel linguistic theme, semantic prosody, has attracted considerable attention in ESL/EFL learning and teaching. The concept of semantic prosody was first introduced to the public in 1993 by Bill Louw, who defined it as “the consistent aura of meaning with which a lexical item is imbued by its collocates” (Louw, 1993: 156-159). Later, as the size of corpora has grown, and tools for extracting semantic prosodies have been developed, semantic prosodies have been addressed much more frequently by linguists (Sinclair, 1991; Louw, 1993; Stubbs, 1995, 1996; Partington, 1998; Hunston, 2002; Wei, 2002, 2006; Xiao & McEnery, 2006).

While most of the existent studies of semantic prosody are confined to the description of native speakers’ English, there has not been sufficient work done on the English as Second Language (ESL). Still less work has been undertaken on the English as Foreign Language (EFL) in China. Therefore, this corpus-based study attempts to bridge the gap and supplement the present studies of semantic prosody. Specifically, the present paper attempts to compare the English writings made by Chinese learners and native speakers with the purpose of finding the differences and/or similarities in terms of semantic prosody, hoping to provide some valuable implications for our EFL teaching and learning.

II. REVIEW OF SEMANTIC STUDY

The term *semantic prosody*, also called discourse or pragmatic prosody (Stubbs, 2001), or semantic associations (Hoey, 2003; Nelson, 2006), was coined by Sinclair (1987), who borrowed Firth’s (1957) notion of *phonological prosody*. Semantic prosody was first introduced to the public by Louw (1993). This term has been widely used by Hunston (2002), Partington (1998, 2004), Stubbs (1995, 2001), Tognini-Bonelli (2001) and Tribble (2000), among many others. Currently semantic prosody has become an important concept in corpus linguistics (Whitsitt, 2005).

In terms of classification of semantic prosody, Stubbs (1996: 176) proposes that some words have a predominantly negative prosody, a few have a positive prosody, and many more words are neutral or mixed in this respect. If the collocates that a node word attracts are mostly of strong negative semantic characteristics, the node word bears a strong negative semantic prosody. If the collocates are mainly positive words, then the node word is endowed with a positive semantic prosody. If both positive and negative collocates exist in the context, the node word can be said to bears a neutral or mixed semantic prosody.

A. Major Relevant Studies Undertaken by Scholars Abroad

In Louw’s view, semantic prosody cannot be retrieved reliably through introspection, and consciously upsetting a word’s semantic prosody can achieve an ironic effect. Louw (2000) further claimed that negative semantic prosodies were much more frequent than positive ones. Another researcher, Partington (1998), defined semantic prosody as “the spreading of connotational coloring beyond single word boundaries” (p. 68). In this definition, semantic prosody is more strongly associated with *connotation*. Stubbs (1995) and Hunston (2002) expanded the notion of semantic prosody by suggesting that in addition to collocating with positive or negative groupings of words, lexical items can also collocate with semantic sets. According to Hunston, “A word may be said to have a particular semantic prosody if it can be shown to co-occur typically with other words that belong to a particular semantic set”(p.104). For example, the word *unemployment* displays a tendency to collocate with the semantic set of *statistics*. Tribble’s definition further broadened the notion of semantic prosody in that a lexical item can be featured with both a global semantic prosody in terms of the whole language, and a local semantic prosody in a certain context or genre. His analysis of the word *experience* in a genre-based corpus showed that there is a local semantic prosody of *experience* in the genre examined.

Xiao and McEnery (2006) explore the semantic prosody of near synonyms from a cross-linguistic perspective and compare the phenomena of semantic prosodies existing in Chinese and English. And the study shows that semantic prosody and semantic preference are as observable in Chinese as they are in English. Remarkably, it is the first attempt to study the semantic prosodic behaviour existing in the Chinese language and one of cross-linguistic investigation of semantic prosody.

B. Major Relevant Studies Undertaken by Scholars in China

In the past few years, more and more Chinese scholars have paid special attention to the phenomenon of semantic prosody and have conducted some insightful studies.

Wei (2002) introduces three approaches to the study of semantic prosody: data-based approach, data-driven approach and integrated approach. The significance of this study lies in providing effective approaches for the further and deeper exploration of semantic prosody.

Pan and Feng (2003) also introduce the concept of semantic prosody, illustrate the exploitation of semantic prosody and discuss its various applications. But this study is far from a systematic and in-depth exploration.

Wang and Wang (2005) examine the semantic prosody of CAUSE. The study shows that great differences exist in the semantic prosody of CAUSE between Chinese learners of English and English native speakers. Chinese learners of English underuse the typical negative semantic prosody and at the same time overuse the atypical positive semantic prosody. However, the study is confined to the semantic prosody of CAUSE without adequate attention to its collocation patterns.

Wei (2006) is a significant contrastive study of the semantic prosodic features of the lexical combinations in the Chinese learners' English writings. The study shows that there exist three major types of word combinations in learner's interlanguage, including typical collocations, interlanguage collocations and unusual collocations. The paper argues that collocational acceptability is no less important than grammatical acceptability in second language learning and.

Considering the fact that the study of semantic prosody in China is still under its preliminary stage, some mistakes or imperfection in those studies are unavoidable. And it is hoped that the present study is able to take advantage of their strength and supplement the current studies.

III. RESEARCH DESIGN

The present study is based on CLEC and BROWN. CLEC is the first learner corpus constructed by Chinese scholars and also the most frequently used one. It is claimed to be reliable not only for the amount of the sampled data, but also for the sampling process. The samples are all original student writings without any correction and are from diverse sources so that the corpus covers learners' written output widely enough. BROWN is a prestigious corpus of English as native language. This study attempts to compare the semantic prosodic features of the lexical item COMMIT in CLEC and BROWN. Here capital letters are deliberately used to refer to a lemma, which stands for all the word-forms of the verb. That means, COMMIT stands for *commit*, *commits*, *committed*, *committing*. As is mentioned above, COMMIT is proved to be a typical example which exhibit obvious semantic prosody in native English. The detailed research procedures are described as follows:

First, the four items are in turn concordanced through Wordsmith 3.0 in CLEC and Brown and all the collocates of the items are extracted. Then with the aid of EXCEL program, the MI value and Z-score are calculated for each extracted collocates. Those with a MI score of 3 or higher and a Z-score of 2 or higher are considered to be significant collocates. After that, the semantic prosodic features of the four items are generalized through analyzing the semantic features of their significant collocates. Finally, the data from the two corpora are compared to find the differences and/or similarities between native speakers and Chinese English learners.

By exploring the semantic prosody of COMMIT in CLEC, the study attempts to address the following two questions:

1. Does Chinese learners' English exhibit semantic prosody as native speakers' English does?
2. How different (or similar) are the semantic prosodic features reflected in the English writing made by native speakers and Chinese learners? Why?

IV. RESULT AND DISCUSSION

A. The Semantic Prosodic Features of COMMIT in Brown

In Brown corpus, we find 51 occurrences of COMMIT. We first observe the 51 concordance lines with eyes and look at the left and right collocates in the span of five. And then we attempt to construct the colligations and find out the collocates which suit the colligation.

TABLE 1
SIGNIFICANT COLLOCATES OF COMMIT IN BROWN

Collocates	F(c)	F(n,c)	MI	Z-score
suicide	17	6	12.777	91.711
adultery	3	1	12.694	36.384
murder	76	5	10.353	36.034
crimes	14	2	11.472	33.651
murderer	21	2	10.887	27.452
sin	54	3	10.109	25.631
theft	10	1	10.957	19.893
perfection	11	1	10.820	18.963
conscience	40	1	8.9577	9.872
assumption	41	1	8.922	9.7482
errors	44	1	8.820	9.403
assault	95	1	7.709	6.316
fraud	112	1	7.472	5.792
robbery	123	1	7.34	5.5110
effort	145	1	7.100	5.047
death	286	1	6.120	3.461

i. commit+N

For this colligation, “commit something” means “perform something negative”, which matches the first dictionary sense. In the result, 26 of the 51 lines have this colligation, accounting for 51%. Typical collocates are *adultery, murder, sin, suicide, crimes, sacrilege, death, theft* and they all refer to something bad or wrong. However we cannot rush to a conclusion that these collocates have statistical value until we are able to put them into statistical measures.

The MI score and Z-score of these collocates are then calculated and all the significant collocates of COMMIT are displayed in Table 1.

In the first column of the table, collocates of COMMIT are listed; for the second and third columns, F(c) refers to the frequency of these collocates in Brown and F (n, c) refers to the frequency of the co-occurrence of COMMIT and its collocates. When its MI score reaches 3 or above and Z-score is 2 or above, the word is considered to be a significant collocate.

As the table indicates, the MI scores and Z-scores of *suicide, adultery, murder, crimes, sin, sacrilege, death, theft* are both far higher than 3, so they are all considered to be significant collocates of COMMIT. Apparently, COMMIT used in this colligation tends to attract negative words and so exhibits an obvious negative semantic prosody.

ii. commit+N+to+N

In this colligation, COMMIT means “to give, entrust, and hand over to, for safekeeping or treatment.” Actually there are only five instances of this colligation in our concordance result. Three are respectively “to commit somebody to Bridewell Prison/House of Detention/Chester Asylum”. Certainly a prison or asylum is not a desirable place. The other two do not show clearly anything bad or wrong. Even when an extended context is obtained and studied, no positive or negative connotation is implied. Therefore, the item COMMIT occurring in this colligation basically shows mixed semantic prosody.

iii. be+committed+to+N

Of 51 concordance lines, 15 fall into this colligation, nearly accounting for 30%. Three lines are cited here as example.

1. Since American life **is committed** above all **to** productivity and more lea
2. war mount in frenzy . The country **is committed to** the doctrine of security by military means
3. force of nature. We are already **committed to** establishing man's supremacy over nature

These lines all entail an obligation or dedication of the agent. The collocates here include *plan, modernization, productivity, supremacy, faith, doctrine*, which seem to refer to something favourable. If we make a careful study of the broader context, we find that most of the instances are not favorable indeed.

iv. commit + oneself + to + N/V-ing/V

As for this colligation, another dictionary, Collins COBUILD dictionary provides clearer explanations than OALD.

“If you commit yourself to something, you say that you will definitely do it. If you commit yourself to someone, you decide that you want to have a long-term relationship with them.”

There are six lines for this colligation in the concordance result, which are displayed below:

1. Fifty-two companies started or **committed themselves to** new plant construction , t
2. Having once **committed himself to** an ideal which he considers worthwhile , he inevitably
3. before he had **committed himself to** a life of austerity and denial ? Now , when everything
4. we do not even need to **commit ourselves to** assessing on the same inconceivable scale
5. assert the value of such study , we **find ourselves committed to** an important assumption .
6. influences still appears as the cardinal sin realtors **see themselves committed** to avoid .

While the collocates *plant construction, to avoid sin* and *a worthwhile ideal* are something favourable or beneficial, *a life of austerity and denial* is something undesirable. *Assumption* and *scale* are neither favourable nor unfavourable.

Therefore half of the instances exhibit a positive semantic prosody.

To sum up, the lexical item COMMIT has a broad range of colligations and varied senses in native speakers' English. On the whole, COMMIT shows a negative semantic prosody especially when it is used in the colligation of "commit+N". When used in other colligations, it demonstrates a mixed semantic prosody, either negative or positive or neutral. Then how is COMMIT used in English written by Chinese learners? It will be elaborated on in next section.

B. The Semantic Prosodic Features of COMMIT in CLEC

In CLEC corpus, 191 instances of COMMIT have been located, which is a much higher frequency, compared with that of Brown. However, when we look at the left and right collates in the concordance lines, we find a rather limited range of vocabulary. *Crime* and *crimes* occur 127 times and *suicide* occurs 25 times and in most cases, the three word-forms serve as the object or subject of COMMIT in CLEC. And then we inquire further to find that most of the concordance lines are extracted from two articles of ST6 (a sub corpus of advanced English majors) of CLEC. The two articles are about "crimes" and "euthanasia". Here dozens of lines are displayed for the reader to understand the situation (see Table 2).

Likewise, the colligations of COMMIT established from CLEC are surprisingly simple. 188 of 191 instances display the colligation of "commit+N", though some appear in the pattern of passive voice. Only three occurrences of "be+committed+to+N" exist. It's evident that Chinese learners are extremely familiar with the colligation of "commit+N" but have little idea of other diversity. Matching this colligation with the frequent collocates, we can see they use "commit+N" (*crime, crimes, suicide, murder, killing, euthanasia*) in a great majority of cases.

Next a corpus-driven approach is employed to get statistical evidence. And all the significant collocates are demonstrated in the following table:

TABLE 2
SIGNIFICANT COLLOCATES OF COMMIT IN CLEC

Collocate	F(c)	F(n,c)	MI	Z-score
suicide	33	25	12.052	145.530
crimes	170	56	10.850	143.414
crime	456	71	9.769	110.700
homicide	2	2	12.452	47.304
act	140	12	8.908	33.603
punishment	188	7	7.705	16.684
murder	133	6	7.911	16.646
killing	388	6	6.438	9.611
case	290	3	5.857	5.390
euthanasia	1241	7	4.982	5.603
evil	252	3	6.060	5.853

As is clearly shown in Table 2, the collocates *suicide, crime* and *crimes* have such a high z-score that is far above the significant level. And the other eight collocates also have a high significance. If we compare Table 1 with Table 2, we can easily see that *crime, crimes, suicide* and *murder* are rather frequent collocates of COMMIT in both corpora. Nonetheless, the frequency of the first three words is much higher in CLEC than that in Brown while the frequency of *murder* in CLEC is much lower than that in Brown.

Apart from *act*, which is quite neutral in connotation, all the significant collocates of COMMIT, *crime, crimes, suicide, homicide, punishment, murder, killing, euthanasia* and *evil* refer apparently to bad or wrong things. That proves that Chinese English learners have mastered two typical collocations of *to commit a crime* and *to commit suicide* well and are quite aware of its negative semantic prosody. It should be noted that they also use a collocation, *to commit a case*, which is strange to native speakers. It might be due to the negative transfer of their native language, where there is a collocation "*fan'an*". So *commit a case* here could be a literal translation of "*fan'an*".

If we compare the data from CLEC with the data from Brown, we can see that both Chinese learners and native speakers are aware of the negative semantic prosody of COMMIT. However, Chinese students merely use the item in the single colligation of "commit+N" and in the narrow sense of "being involved in crime" and have the least idea of the other senses in other colligations, while native speakers command a much broader range of colligations and more varied senses. And their English exhibits a full-scale semantic prosody for *commit*, not a simply negative semantic prosody.

This result suggests that Chinese learners have neither a complete idea of the collocational patterning nor a complete profile of the semantic prosody of *commit*. That may result from several factors. One factor is Grammar Translation teaching method and so the teacher. Although communicative language teaching approach has been claimed to be implemented in China for twenty years, Grammar Translation method is still prevalent in most English classrooms of Chinese high schools and colleges. And knowledge-based teaching concept is still hidden in many English teachers' mind. Influenced by this method and concept, the teachers tend to offer the translation equivalent of a new word in vocabulary instruction. Specifically for the word *commit*, they may tell students that "*commit* means *to do or perform*, equivalent to *gan* or *zuo* in Chinese". To deal with this situation, the traditional method of vocabulary instruction must

be improved and the traditional teaching concept should be changed.

In conclusion, there are differences as well as similarities in terms of collocation and semantic prosody between the English writing made by native speakers and Chinese learners. In some cases, learner English does exhibit similar semantic prosody. In other cases, Chinese learners tend to collocate a node word showing a positive prosody with collocates that show negative semantic characteristics, or vice versa. In addition, from the perspective of collocation, distinct differences exist between the English writing made by native speakers and Chinese learners, even when the two show similar features of semantic prosody. For one thing, Chinese learners overuse a narrow range of general words as collocates and underuse variety of specific words. For another, Chinese learners use a number of interlanguage collocates and unusual collocates, which makes their English sound quite unnatural and less idiomatic.

V. CONCLUSION

The present corpus-based study has made a comparison of the semantic prosodic features of COMMIT between CLEC and Brown. The results indicate that there are differences as well as similarities between English writings by native speakers and Chinese learners in terms of collocation and semantic prosody. Based on these findings, the study can provide significant implications for EFL teaching and learning, especially for vocabulary instruction as well as dictionary compilation. Firstly, it is suggested that the knowledge of semantic prosody should be transferred to students so that they are able to gain due awareness of semantic prosody. Secondly, it implies that the knowledge of semantic prosody can also provide insight into the teaching of near synonyms. Near synonyms with identical or similar denotational meaning can be distinguished in their collocational behaviours and semantic prosodies. Thirdly, the deficiency in typical collocations in learner English reveals that more emphasis should be put on the teaching of collocation instead of teaching separate words without context. Lastly, given most contemporary English learner dictionaries neglect the explanation of semantic prosody, it is advisable to compile dictionaries that provide adequate information concerning semantic prosody features of vocabulary entries for English learners.

REFERENCES

- [1] Firth, J. (1957). *Papers in linguistics, 1934-1951*. London: Oxford University Press.
- [2] Hoey, M. (2003). Lexical priming and the qualities of text. Retrieved October 14, 2008, from: <http://www.monabaker.com/tsresources/LexicalPrimingandthePropertiesofText.htm>
- [3] Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge, UK: Cambridge University Press.
- [4] Louw, B. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair* (pp. 157-176). Amsterdam: John Benjamins.
- [5] Louw, B. (2000). Contextual prosodic theory: Bringing semantic prosodies to life. In C. Heffer, H. Sauntson, & G. Fox (Eds.), *Words in context: A tribute to John Sinclair on his retirement* (pp.48-94). Birmingham, UK: University of Birmingham.
- [6] Nelson, M. (2006). Semantic associations in Business English: A corpus-based analysis. *English for Specific Purposes*, 25, 217-234.
- [7] Pan, P., & Feng, Y. (2003). Corpus-based analysis of semantic prosody and its applications. *Contemporary Linguistics*, 5(4), 359-366.
- [8] Partington, A. (1998). *Patterns and meanings: Using corpora for English language research and teaching*. Philadelphia, PA: John Benjamins.
- [9] Partington, A. (2004). "Utterly content in each other's company": Semantic prosody and semantic preference. *International Journal of Corpus Linguistics*, 9(1), 131-156.
- [10] Sinclair, J. (1987). *Looking up: An account of the COBUILD project in lexical computing and the development of the Collins COBUILD English language dictionary*. London/Glasgow: Collins.
- [11] Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford, UK: Oxford University Press.
- [12] Stubbs, M. (1995). Collocations and semantic profiles: On the cause of the trouble with quantitative methods. *Function of Language* 2/1: 1-33.
- [13] Stubbs, M. (1996). *Text and Corpus Linguistics*. Oxford: Blackwell.
- [14] Stubbs, M. (2001). *Words and Phrases*. Oxford: Blackwell.
- [15] Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam: John Benjamins.
- [16] Tribble, C. (2000). Genres, keywords, teaching: Towards a pedagogic account of the language of project proposals. In L. Burnard & T. McEnery, (Eds.), *Rethinking language pedagogy from a corpus perspective: Papers from the Third International Conference on Teaching and Language Corpora* (pp. 74-90). New York: Peter Lang.
- [17] Wang, H., & Wang, T. (2005). A contrastive study on the semantic prosody of CAUSE. *Modern Foreign Language*, 28(3), 297-307.
- [18] Wei, N.X. (2002). A corpus-driven study of semantic prosodies in specialized texts. *Modern Foreign Language*, 25(2), 165-175.
- [19] Wei, N.X. (2006). A corpus-based contrastive study of semantic prosodies in learner English. *Foreign Language Research*, 132, 50-54.
- [20] Whitsitt, S. (2005). A critique of the concept of semantic prosody. *International Journal of Corpus Linguistics*, 10 (3), 283-305.
- [21] Xiao, R. & A. McEnery. (2006). Collocation, semantic prosody, and near synonymy: A cross-linguistic perspective. *Applied Linguistics*. 27/1: 103-129.

Changhu Zhang was born in Dezhou, China in 1974. He received his M.A. degree in linguistics from Shandong University, China in 2007.

He is currently a lecturer in the Department of Foreign Languages, Dezhou University, Dezhou, China. His research interests include second language acquisition and corpus linguistics.