Journal of Arts & Humanities

Volume 08, Issue 01, 2019: 01-15
Article Received: 11-12-2018
Accepted: 08-01-2019
Available Online: 13-01-2019
ISSN: 2167-9045 (Print), 2167-9053 (Online)
DOI: http://dx.doi.org/10.18533/journal.v8i1.1552

A Corpus-based Contrastive Study on Semantic Prosody of English near Synonyms: A Case Study of Motive and Motivation

Qian WANG[1]

ABSTRACT

Corpus-based studies on semantic prosody of synonyms provide a new lens to teaching and learning synonyms in English. However, using corpus-based approach to facilitate teaching and learning English synonyms is still in its infancy in China. This study, based on Chinese Learner English Corpus (CLEC) and Corpus of Contemporary American English (COCA), conducts a contrastive analysis of English near synonyms "motive" and "motivation" in terms of frequency, collocation features and semantic prosody between Chinese English Foreign Language learners and native English speakers. Findings suggest that the node word is underused by English learners. Furthermore, motive tends to co-occur with collocates related to illegal activities while motivation is apt to appear with collocates involved with normal or beneficial activities. In addition, Chinese language Learners overuse negative semantic prosody of motivation and positive semantic prosody of motive. Reasons for underuse and misuse are explored and research implications are discussed as well.

Keywords: Collocation, Corpus, Semantic Prosody, Synonyms.

1. Introduction

It is well known that English has a large number of loanwords or borrowed words from other languages. As reported, synonyms are abundant in English, accounting for 60% of vocabularies (He, 2003). Hence, to have a good command of vocabularies to some extent, lies in a better grasp of synonyms for language learners.

The fact that synonyms in English constitute the complicated system of English vocabulary, in deed, poses barriers to Chinese EFL (English Foreign Language) learners in terms of lexical accuracy for expression. Language learners, when exposed to a list of near synonyms to choose from, tend to resort to their intuitions for distinguishing subtle semantic difference between words. As a result, lexical and pragmatic misuses of synonyms arising from ignorance of semantic prosody commonly occur among Chinese EFL learners (Wei, 2006; Xiao & McEnery, 2006).

[1] School of Foreign Studies, Northwestern Polytechnical University, Faculty of Humanities, Hongkong Polytechnic University. E-mail: wangq@nwpu.edu.cn

Collocation and semantic prosody, as an emerging trend of lexicology research, could address the challenge of discriminating near synonyms in English vocabularies with assistance of corpora. Semantic prosody, as an important feature of collocation, can reveal "the speaker or writer's attitude or stance towards, viewpoint or feelings about the entities and propositions that he or she is talking about" (Hunston & Thomsom, 2000, p. 5). In addition, semantic prosody is considered as a primary factor in the choices of utterance and has a determining role in the co-selecting mechanism of lexical grammar (Sinclair, 1996; Stubbs, 2009). Hence, semantic prosody opens up a new field for lexical semantics and lexicography, which provides a new lens to observe, analyze and describe the lexical behavior. Without doubt, semantic prosody could, to a large extent, equip learners with language idiomaticity, thereby facilitating effective English communication. However, this concept is still new to a large number of EFL learners who fail to realize the need and urgency of grasping semantic prosody of lexical items.

These two decades have witnessed a substantial body of studies devoted to semantic prosody, as demonstrated by the contributions of, for example, Sinclair (1991,1996) and Stubbs (1995, 1996) in monolingual studies, as well as Partington (1998, 2004), Sardinha (2000), Tognini-Bonelli (2001) and Xiao and McEnery (2006) in cross-lingual studies. A few researchers in China also endeavored to investigate the different use of synonyms of English learners (Huang, 2007; Lu, 2010; Sun, 2017), but contrastive analysis between EFL learners and native speakers of English synonyms is still under-explored. Moreover, in these studies, either verbs or adjectives were addressed, leaving much room to explore the characteristics of prosodic use of nouns among Chinese EFL learners, misuses in particular for bringing about more pedagogical implications.

In light of this, this study, adopting corpus-based approach, conducts a contrastive analysis of near synonyms in terms of frequency, collocation features and semantic prosody between Chinese EFL and native English speakers. A pair of easily-confused synonyms for learners "motive" and "motivation" is selected and examined based on Chinese learner English corpus (CLEC) and Corpus of Contemporary American English (COCA) to address the following questions:

1) What are the differences of the synonyms in terms of frequency between Chinese EFL learners and native English speakers?

2) What are the differences of the synonyms in terms of collocation features between Chinese EFL learners and native English speakers?

3) What are the differences of the synonyms in terms of semantic prosody between Chinese EFL learners and native English speakers?

2. Literature review

2.1 Definition of some Key Terms

2.1.1 Near Synonyms

In *Longman Dictionary of Contemporary English*, synonym is defined as "a word with the same or nearly the same meaning as another word in the same language". Synonyms can be grouped into absolute synonyms and near synonyms. The former refers to words that can be used to replace each other in any circumstance like *truck* and *lorry* whereas the latter relates to a series of words with "very similar cognitive or denotative meanings but distinct collocational or prosodic features" (Partington, 1998, p. 77) such as *cause, bring about, lead to.* In this sense, near synonymous words cannot be collocationally interchanged.

2.1.2 Collocation

Scholars' definitions of collocation are many-folds, ranging from lexical to grammatical perspectives, and from descriptive to prescriptive forms. The linguistic term, collocation, is originally put forward by Firth, who states "You shall know a word by its company it keeps" (Firth, 1957, p. 14). According to Firth, collocations of a given word are statements of the habitual or customary places of that word. Kjellmer later brings grammar elements into his definition: "a collocation is a sequence of words that occur more than once in identical forms and which is grammatically well-structured" (Kjellmer, 1987, p. 33). Hoey perceives that the premise of lexical items to be counted as collocations is

when they appear together "with greater than random probability in its (textual) context" (Hoey, 1991, p. 6). Sinclair (1991) claims that collocation is the co-occurrence of two or more adjacent words in a text. When describing collocation, he proposes a set of related concepts such as node, collocate and span. According to Sinclair, node means the word being studied, collocate refers to any word that occurs in the specified environment of node and such a specified environment is called span.

2.1.3 Colligation

Colligation describes grammatical patterns of words, which refers to "the inter-relation of grammatical categories in syntactical structure" (Firth, 1957, p. 165). Hoey contends that "colligation can be defined as the grammatical company a word keeps and the positions it prefers" (Hoey, 2000, p. 234). Sinclair (2004) maintains that colligation is "the co-occurrence of grammatical choices" (Sinclair, 2004, p. 142), which is closely related to collocation, but is "more abstract than collocation" (Sinclair, 1996, p. 103).

Therefore, both collocation and Colligation are types of phraseologism (Gries, 2008). However, colligation is involved in the relationship among grammatical classes, while collocation focuses on the words which belong to these grammatical classes. Generally speaking, collocation is taken for the concrete realization of certain colligations. Different as they are, collocation and colligation are entwined and in the meanwhile, cooperative to manifest the deep-seated features of linguistic patterns.

2.1.4 Semantic preference

In Sinclair's extended units of meaning, semantic preference means a lexical set of frequently occurring collocates, which share a semantic feature (Zethsen, 2006). Stubbs perceives semantic preference as "the relation, not between individual words, but between a lemma or word form and a set of semantically related words" (Stubbs, 2001, p. 65).

Semantic preference and semantic prosody are characterized with being distinctive yet interdependent concepts, as Sinclair (1996) argues that "Prosody is at a further stage of abstraction than preference" (Sinclair, 1996, p. 87). Partington is with the view that "Semantic preference is a narrower phenomenon relating the node item to another item from a particular semantic set than prosody which can affect wider stretches of text" (Partington, 2004, p. 151).

2.1.5 Semantic prosody

Definitions of semantic prosody are available in a number of versions. Louw (1993), first defines it as "a consistent aura of meaning with which a form is imbued by its collocates" (Louw, 1993, p.157), and then calls for a revised working definition as follows: "A semantic prosody refers to a form of meaning which is established through the proximity of a consistent series of collocates, often characterizable as positive and negative, and whose primary function is the expression of the attitude of its speaker or writer towards some pragmatic situation" (Louw, 2000, p. 57). This definition argues that semantic prosody rests in collocates rather than the node word itself and more importantly, semantic prosody goes beyond a lexico-grammatical level by extending to both semantics and pragmatics. Sinclaire (2000), who describes semantic prosody from the pragmatic perspective, is also convinced that semantic prosody explains the reason for the selection of an certain lexical item instead of another. Another researcher, Stubbs (1996), defines semantic prosody as "words occur in characteristic collocations, showing the associations and connotations they have and therefore the assumptions which they embody" (Stubbs, 1996, p. 135). In a sense, he relates semantic prosody to the practical use of co-occurring lexis and emphasizes the interplay between semantic prosody and collocation. Partington (1998) considers semantic prosody as "the spreading of connotative coloring more beyond strongly single boundaries" (Partington, 1998, p. 68). What is more, Partington holds that context plays an essential role in determining the semantic and the attitudinal meaning of words. Hunston &Thompson (2000) assume semantic prosody embodies "speakers' and attitudes, feelings and stances about facts in communications" (Hunston &Thompson, 2000, p. 5).

In a nutshell, though semantic prosody is defined by researchers from different perspectives, the common features of semantic prosody have been agreed on. For example, semantic prosody should be built up by collocates it frequently co-occurs with in contexts and it is not only affiliated to

lexico-grammatical level but also to pragmatic level.

In relation to the question of classifying semantic prosody, this study adopts the one proposed by Stubbs (1996) who divides semantic prosody into three categories: positive semantic prosody, neutral semantic prosody and negative semantic prosody. In positive semantic prosody, almost every word around the node word bears affectively positive meanings and semantic atmosphere, while in negative semantic prosody, the node words attract collocates with unfavorable meanings, making the context covered with an unpleasant atmosphere. In neutral semantic prosody, collocates of the node word show no tendentiousness for any affective meaning.

2.2 Previous studies on semantic prosody

The first area being investigated is monolingual studies of semantic prosody. Sinclair (1991, 1996) first scrutinizes the collocates following *break out* and *set in* within context, suggesting that the two phrases are often associated with negative prosody. Louw (1993) explores *build up* and notices that its semantic prosody changes when used as a transitive verb or intransitive verb.

Stubbs (1995) analyzes *cause* and *provide* from the COBULD Corpus and finds that 90% of lexical items that collocate with *cause* are unfavorable; on the contrary, *provide* tends to co-occur with pleasant lexical items. Channell (2000) studies 1020 examples of the verb *roam,* demonstrating that *roam* becomes negative when associated with "street" though *roam* is generally neutral in dictionary. Partington (2004) examines a set of synonyms of *happen* and claims that *set in* had the worst prosody, followed by *happen*; *take place* also tends to occur in unpleasant environment while *come about* does not evince any particular inclination.

Another line of research conducted is on semantic prosody studies across different languages. Partington (1998) carries out a comparative study of semantic prosody between near synonyms in English and Italian and argues that though English word *impressive* and the Italian word *impressionante* are similar in spelling and meaning, *impressive* bears a positive prosody while *impressionante* exhibits negative semantic prosody. Sardinha (2000) also makes a comparative analysis of semantic prosody between English and Portuguese, suggesting that the semantic prosody of some equivalent words are either the same or different.

Therefore, he lays stress on the significance of a good command of knowledge related with corpus-based semantic prosody for translators and bilingual dictionary editors. Tognini-Bonelli (2001) also investigates the semantic prosody of *in case of* in English and its Italian equivalent and concludes that "semantic prosodies are often linked to language varieties and restricted the varieties" (Tognini-Bonelli, 2001, p.138). Xiao and McEnery (2006) conduct a comparative study on the semantic prosody of synonyms between Chinese and English, which illustrates that the Chinese equivalents are "more sharply divided between the clearly negative and positive ends of the continuum" (Xiao &McEnery, 2006, p. 120).

3. Research methodology

3.1 Research corpora

The corpora used in the current study are Corpus of Contemporary American English, known as COCA and Chinese Learner English Corpus known as CLEC.

Created by Mark Davies of Brigham Young University, COCA is the largest freely-available corpus of English. The statistical data in COCA provide insights into varieties of English, encompassing approximately 570 million (20 million words each year from 1990 to 2017) words, sampled from spoken, novel, popular magazines, newspapers, and academic texts. The distributions of sub-corpora in COCA are listed below. The data from the spoken genre will be excluded to be in parallel with CLEC for comparison, so the tokens total around 450 million words.

Table 1: Distributions of COCA

Sub-corpora in COCA	Spoken	Fiction	Popular magazines	Newspapers	Academic journals	Total
Words (million)	119	112	117	114	113	570

Known as one of the most representative English learner corpora, and developed by Gui & Yang in China in 2008, CLEC consists of 5 sub-corpora of the samples collected from English learners' original writings from diverse sources such as in classrooms or national standardized English tests in China. It represents different levels of proficiency of Chinese EFL learners ranging from high school students to English majors in university.

Table 2: Distributions of CELC

Sub Corpora and Levels of Proficiency	Words	Total
ST2- High School Students	208,088	
ST3- Non-English Major Students in Freshmen and Sophomore Year	209,043	
ST4- Non-English Major Students in Junior and Senior Year	212,855	
ST5- English Major Students in Freshmen and Sophomore Year	214,510	1,070,602
ST6- English Major Students in Junior and Senior Year	226,106	

3.2 Node words

The synonymous pairs "motive" and "motivation", "rather" and "fairly" are not chosen randomly. Firstly, the first pair shares the meaning "something that causes a person to act and the second pair shares the meaning "to some degree" as defined by Merriam-Webster Thesaurus Online, which indeed confuse a majority of English learners in particular. Secondly, these pairs of synonyms satisfy the need of more than 50 concordance lines comprising node words (Wei, 2002) to analyse their semantic features since the two corpora demonstrate high frequency counts of these node words.

3.3 Research tools

Developed by the Japanese professor Laurence Anthony in 2002, AntConc could provide KWIC (key words in context), frequency, concordance lines and collocates as well as calculate MI score and extract lexical chunks.

By employing the research tool AntConc3.2.1w (2014), frequency of node words, collocates of node words, concordance lines and MI score in two corpora could be identified.

3.4 Statistical measures

MI score (Mutual Information) score, indexes mutual attraction between a node word and its collocations, indicating different strength of the collocation between the lexical items. Simply put, high MI score demonstrates that two words are inclined to co-occur with each other more frequently; low MI score means two words do not depend on each other and co-occur in accident whereas negative MI score reveals the two words never co-occur with each other.

As contended by Hunson (2002), the collocate is considered as significant one if its MI score is equal or higher than three.

Additionally, Log-Likelihood (LL) Calculator developed by Xu (2003) From Beijing International Studies University will be utilized to verify whether the difference concerning frequency of the node word in two corpora is significant or not. The significance level of LL value is corresponding to P value. On the premise that the degree of freedom is 1, if LL value is higher than the critical values, namely, 3.84, 6.64 and 10.83, the LL value is significant. When LL value is higher than 3.84 and P value is lower than 0.05, the difference is significant (95% sure). When LL value is higher than 6.63 and P value is lower than 0.01, the difference is very significant (99% sure). When LL value is higher than 10.83 and P value is lower than 0.001, the difference is extremely significant (99.9% sure).

3.5 Research procedures

First, the frequency of each node word in two corpora is calculated. The word frequency per million is employed given the different capacities of the two corpora. Meanwhile, log-likelihood value is used to determine if the difference in frequency between the two corpora is significant.

The next step involves with investigation of collocation features. The extract of concordance lines of each node word will be set with span of ±5 for *motive* and *motivation*. Afterwards, the top 100 most frequently used significant collocates are extracted for analysis of semantic prosody of each node

word. Only significant collocates (MI≥3) are under consideration. Then, the selected collocates are classified into different semantic sets, ready to be analyzed.

Finally, the semantic prosodies of node words are examined with manual annotation within context in COCA and CLEC, which will be followed with explorations of possible misuses with illustrations on part of Chinese EFL learners.

4. Findings and discussions

4.1 A contrastive analysis of frequency of node words

4.1.1 Frequency of motive and motivation in COCA and CLEC

The calculation of word frequency in one million words is employed to detect the usage rate of synonyms due to the fact that COCA has approximately 0.4 billion tokens whereas CLEC consists of about 1 million words. In addition, log-likelihood (LL) value is taken into account to ascertain whether the differences on use frequency of synonyms between Chinese EFL learners and English native speakers are significant or not.

In this study, concordance lines of *motive* used as nouns are considered in both corpora in comparison with *motivation* which is used only as a noun. The frequency of *motive* and *motivation* in CELC and COCA are presented as follows.

Table 3: Frequency distribution of *motive* and *motivation*

Synonyms	CLEC		COCA		LL Value
	Frequency	Per million	Frequency	Per million	
motive	32	8.6	9956	18.65	-25.03
motivation	74	19.9	15325	28.7	-11.17

As shown in Table 3, the frequency of *motivation* in COCA is 1.5 times more than *motive*, suggesting the wider use of *motivation* in native English speaker corpus. By comparison, the frequencies of *motive* and *motivation* in CLEC exhibit similar trend with that in COCA, demonstrating that Chinese EFL learners are inclined to employ *motivation* other than *motive* in their expressions as well.

It is noteworthy that the corresponding LL value is -25.03 for *motive* and -11.17 for *motivation,* and the two negative LL values indicate that Chinese EFL learners tend to underuse *motive* and *motivation,* compared with native English speakers.

4.2 A Contrastive analysis of collocation features of node words

4.2.1 Collocation features motive and motivation

As a matter of fact, enhancing Chinese EFL learners' competence of employing proper collocations has always been a pressing concern in language teaching. Having a good command of collocation features of a lexical item, in essence, would bring about immense benefits in terms of helping learners to locate idiomatic and typical collocations in English and in the meanwhile, minimizing their unusual or irregular collocations.

The top 100 significant collocates (MI≧3, Frequency>10) of synonyms *motive* and *motivation* in COCA and CLEC are identified and extracted for analysis. Here, only the first 20 significant collocates with higher MI score are presented because of the space limit.

In line with a number of concordances randomly retrieved from two corpora, it could be recognized that the two words share colligations such as N+V, V+N, and ADJ+N, N+PRE, PRE+N. For analyzing collocational behaviors of the two words, all colligations within the range of -5/+5 are located to identify significant collocates in COCA and CLEC.

4.2.1.1 Collocation features of *motive* and *motivation* in COCA

20 top significant collocates of *motive* and *motivation* in COCA are shown as below.

Table 4: Significant collocates of *motive* and *motivation* in COCA

COCA			
motive		**motivation**	
Collocate	**MI Score**	**Collocate**	**MI Score**
Ulterior	12.16	Self-determined	9.87
Impugning	9.8	Extrinsic	9.75
Impugn	9.67	Intrinsic	9.71
Ascribing	8.62	Achievement	6.54
Leisure-time	8.43	Cognition	6.19
Altruistic	8.04	Self-efficacy	6.03
Self-interested	7.06	Orientations	5.89
Autonomous	6.63	Engagement	5.87
Profit	6.5	Self-determination	5.7
Selfish	6.26	Competence	5.37
Extrinsic	6.15	Mastery	5.32
Sinister	6.14	Enjoyment	5.28
Gratification	5.89	Self-confidence	5.28
Spectators	5.88	Motivational	5.23
Robbery	5.76	Academic	5.22
Melodic	5.61	Persistence	5.15
Questioning	5.48	Efficacy	5.1
Questioned	5.43	Avoidance	5.06
Participatory	5.41	Rewards	4.93
Self-interest	5.35	Autonomous	4.92

In COCA, 9 shared significant collocates of *motive* and *motivation* out of top 100 collocates contribute to semantic preferences of "trait" and "normal activity". These 9 shared words include *task, participation, orientation, sport, primary, autonomous, intrinsic, extrinsic,* and *lack*, all reflecting neutral attitudinal meanings. However, there exist collocate differences between *motive* and *motivation.* The word *motive* is inclined to appear with negative descriptive words, such as *murder, ulterior, selfish, underlying, suspicious* connoting "illegal or improper behaviors and desires" whereas *motivation* is apt to emerge with neutral descriptive words, such as *intrinsic, academic, extrinsic, learning*, as well as with words implying "elements beneficial to make progress" such as *understanding, skills, self-determined, confident, creativity*, etc. In addition, a small number of collocates that co-select with *motivation* indicating "changes of situation", covering words like *differences, scales, levels, increase, affect, decrease*, etc could also be observed. In a nutshell, *motive* more often than not co-occurs with words implying "illegal acts" and "negative emotions" while *motivation* is likely to go together with words connoting "beneficial elements for good results or outcomes".

4.2.1.2 Collocation features of *motive* and *motivation* in CLEC

In total, there appear 20 significant collocates of *motive* and 31 significant collocates of *motivation* in CLEC. The 20 most significant collocates are shown in the following table.

Table 5: Significant collocates of *motive* and *motivation* in CLEC

CLEC			
motive		**motivation**	
Collocate	**MI Score**	**Collocate**	**MI Score**
Ulterior	17.14	Energetic	13.6
Unkind	15.56	Suicide	12.95
Violate	12.97	Commit	12.57
Promoting	12.44	Competing	12.49
Pure	11.92	Criminal	12.32
Selfish	11.12	Discrimination	12.16
Criminals	9.38	Vital	11.79
Individual	8.99	Competition	11.68
Competition	8.93	Struggle	11.61
Development	8.87	Stress	11.42
Various	8.39	Pressure	11.35
Killing	8.39	Mercy	11.2
Opportunity	8.09	Lack	10.86
Different	6.58	Improvement	10.72
Improve	6.3	Compete	10.72
Money	5.53	Provide	10.5
Student	5.51	Survive	10.48
World	5.48	Interesting	10.35
Good	4.88	Support	9.89
Society	4.19	Development	9.89

Motive and *motivation* share the most frequently used collocates of the word *competition* in CLEC. It can be seen that *motive* not only co-occurs with *criminals, violate and killing*, but also collocates with words like *development* or *improve* or *opportunity* denoting elements that are conducive to making progress. With reference to the word *motivation*, it both emerges with words indicating "making progress" such as *improvement* and *development* and co-occurs with words involving with "illegal or irritating behaviors" like *commit, suicide, criminal,* or *discrimination.*

4.2.1.3 A Contrastive Analysis of *motive* and *motivation* on Collocation Features between COCA and CLEC

Examing similarities and differences of *motive* and *motivation* existing between Chinese EFL learners and native English speakers, we have several findings. On one hand, *pure, killing, opportunity* are shared significant collocates of *motive* while *lack* and *different* are shared significant collocates of *motivation* between Chinese EFL learner corpus and native English corpus. But, compared with native speakers, Chinese learners demonstrate lack of variability and accuracy of applications of these two words. On one hand, the restricted writing topics may set limits on their choice of words. On the other hand, it can be seen that *motive* and *motivation* share semantic sets in Chinese EFL learner corpus, which is in conflict with what has been shown in native English speaker corpus. For example, *criminal* and *killing* fall into semantic preferences of *motive* in COCA but are also detected as significant collocates of *motive* and *motivation* in CLEC. This, without fail, indicates Chinese EFL learners fail to recognize that the two words share distinctive semantic preference, thus treating them as synonyms that can be used interchangeably. Furthermore, shared collocates with the highest frequency for *motivation* and *motive* such as *competition* and *improvement* in CLEC do not emerge in top 100 significant collocates in COCA, and semantic set of *motivation* denoting "change of situation" never occurs in CLEC. This could possibly suggest that Chinese EFL learners tend to resort to their intuition for patterns of collocation, which lead to their confusion and misuse of lexical items.

4.3 A contrastive analysis of semantic prosody of node words

4.3.1 Semantic prosody of *motive and motivation* in COCA and CLEC

The top 100 significant collocates (MI≥3, Frequency >10) of node words in COCA and CLEC are searched and classified for analysis to determine their semantic prosody.

4.3.1.1 Semantic prosody of *motive* in COCA

The 100 most frequently used significant lexical items (MI≥3) that co-occur with *motive* are listed as follows.

Positive (25)	*autonomous, pure, altruistic, excitement, apparent, affiliation, purely, noble, motivation, participating, communicate, rhetoric, rational, leisure-time, canon, participatory, passion, endorsed, constructs, melodic, leisure, identities, participation, opportunity*
Negative (52)	*ulterior, behind, profit, murder, theft, mistrust, partisan, kill, questioning, questioned, killing, suspect, mixed, selfish, suspicious, drinking, prosecution, robbery, revenge, competitive, spectators, hidden, prosecutors, sinister, wanting, alleged, impugn, desires, suspicious, unclear, suspects, suspected greedy, skeptical, loaded, suspicious, killings, murders, self-interest, jealousy, cynical, obscure, clues, selfishness, impugning, hatred, speculate, speculation, self-interested, gunman, distrust, implicit*
Neutral (23)	*intentions, female, investigators, correlated, rhythmic, underlying, primary, sport, actions, driven, attribute, instructors, strongest, intrinsic, possible, phrases, extrinsic, unconscious, engaging, contexts, ascribing, analyzing, orientations*

Table 6: The semantic prosodic distribution of the collocates for *motive* in COCA

	Motive		
Frequency in Total			100
Semantic prosody	Positive	Negative	Neutral
Frequency counts	25	52	23
Percentage	25%	52%	23%

Out of the 100 occurrences of *motive*, 52 of them show unpleasant emotional coloring. Collocates presenting positive and neutral meaning account for 25% and 23% respectively. Thus, *motive* tends to occur more frequently with negative lexical items. The most frequently used collocates are *ulterior (286), behind (253), profit (246), possible (180), murder (160), primary (96), participation (91), opportunity (91), participating (85), kill (74)*. Their concordances are extracted as follows.

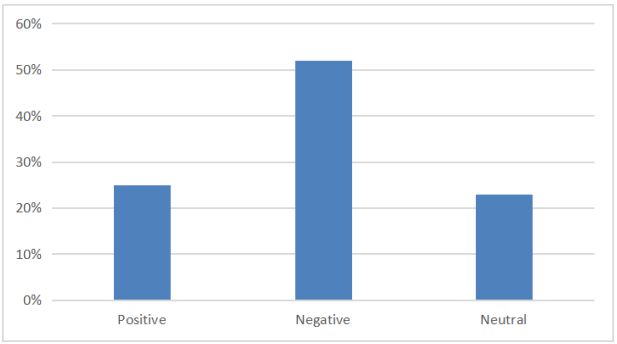

Figure 1: The Semantic Prosodic Distribution of the Collocates for motive in COCA

1. My disappointment lies in those people who somehow think that this change reflects some **ulterior motive**.
2. What is government's official position on the **motive behind** these attacks?
3. Certainly, there has always been a **profit motive** in medicine, but financial considerations have never been as prominent as they are today.
4. Just weeks into the investigation, cops discovered a **possible motive**: money.
5. He doesn't want the jury to think that love was Travis's **motive** for **murder**.
6. A majority of these studies has indicated that the **primary motive** for individuals' use of Facebook is to maintain relationships.
7. This allows educators and practitioners to better understand the **motives** behind students' **participation**.
8. Narrow down who would have **motive**, **opportunity**, and ability to kill her.
9. This study examines various **motives** for **participating** in the sport of cycling.
10. She had a giant financial **motive** to **kill** him and she wanted out.

In the above 10 lines, line 1, 2,3,4,5,8,10 obviously present unfavorable affective meanings which can be detected from the context. Collocates like *behind, profit, possible* and *opportunity* connote positive or neutral implications for literal meanings, but when interpreted with the help of contexts, these 4 words all reveal unpleasant preferences, such as motive behind attack, money for criminal behaviors, or motive and opportunity to kill someone. Line 6 reveals no attitudinal preference and Line 7 and 9 indicate positive affective meanings.

4.3.1.2 Semantic prosody of *motivation* in COCA

Similarly, 100 most frequent significant collocates of *motivation* in COCA are listed as follows.

Positive (49)	*achievement, learning, engagement, increase, skill, primary, goals, understanding, confidence, strategies, competence, enhance, satisfaction, commitment, self-efficiency, perceived, participate, perceptions, creativity, efficiency, beliefs, achieve, self-determined, rewards, engage, expectation, positively, instructional, mastery, feedback, self-determination, enhanced, persistence, enjoyment, succeed, self-esteem, reward, enhancing, attendance, abilities, motivational, inventory, construct, incentive, increased, self-confidence, morale, comprehension, successful*
Negative (6)	*lack, decrease, avoidance, affects, unconscious, affect*
Neutral (45)	*intrinsic, student, academic, extrinsic, levels, sport, behavior, theory, factors, related, differences, participation, influence, exercise, classroom, task, internal, cognitive, psychological, attitudes, fan, theories, variable, athletes, behaviors, external, personality, adolescents, psychology, measured, scales, orientation, emotion, cognition, dimensions, emotions, belongings, desires, autonomous, interpersonal, influences, intentions, contexts, underlying, and influenced*

Table 7: The semantic prosodic distribution of the collocates for *motivation* in COCA

	Motivation		
Frequency in total			100
Semantic prosody	Positive	Negative	Neutral
Frequency counts	49	6	45
Percentage	49%	6%	45%

It is apparent that *motivation* is inclined to co-select both with positive and neutral affective meanings, which constitute 49% and 45% respectively. Only 6% of significant collocates display unfavorable affective meanings.

There are 10 most frequently used significant collocates including *intrinsic (901), achievement (614), student (535), academic (464), learning (375), lack (269), levels (239), engagement (200), behavior (178), skill(169).*

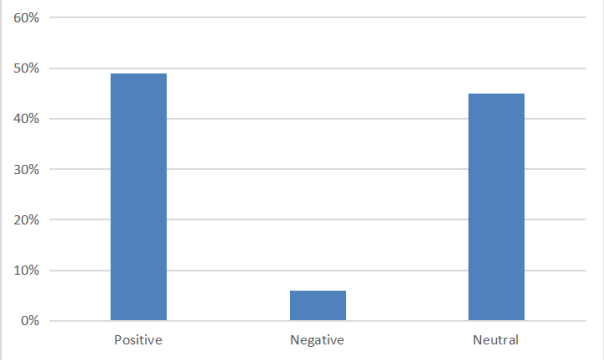

Figure 2: The semantic prosodic distribution of the collocates for motivation in COCA

The concordance lines are listed as follows.

1. Self-actualization contributes to **intrinsic motivation** of what is already internal.

2. Learning experiences, compared with competitive and individualistic ones, promote higher **achievement**, greater **motivation**, and more positive attitudes toward the subject matter and the teacher.

3. Awareness of these three areas of concern promoted continual **student** engagement, **motivation**, and an excellent end product.

4. Supports from an instructor significantly predict college students **academic** emotions, **motivation**, and ultimately their performance.

5. Increasing the authenticity of an assessment is expected to have a positive influence on student **learning** and **motivation**?

6. These students, despite the language barrier, did not **lack** either engagement or **motivation** when it came to discussing and analyzing topics

7. She has noticed that her students' **levels** of self-determination and their **motivation** to write have increased.

8. Increased **motivation** and **engagement** were two often-mentioned benefits for using technology.

9. Some people may receive benefits, even if their **motivation** and **behavior** run counter to the spirit of wellness programs.

10. The most important information needed for such adjustments, such as pretreatment cognitive **skills** and **motivation**, is not available.

It is noticed that positive semantic atmospheres are built up in all concordances above. Although *lack* in line 6 denotes negative implications, the context bestows *motivation* with a good trait via the double negation that indicates favorable affective meaning.

4.3.1.3 Semantic prosody of *motive* in CLEC

In CLEC, only 20 significant collocates of *motive* emerge, among which, 7 collocates convey positive meaning, 7 exhibit neutral sense, and 6 suggest negative affective meanings.

Positive (7)	*competition, development, pure, opportunity, improve, promoting and good*
Negative (6)	*selfish, violate, unkind, ulterior, criminals, killing*
Neutral (7)	*individual, different, various, money, world, student, society*

Table 8: The semantic prosodic distribution of the collocates for *motive* in CLEC

	Motive		
Frequency in total			20
Semantic prosody	Positive	Negative	Neutral
Frequency counts	7	6	7
Percentage	35%	30%	35%

As revealed in Figure 3, in CLEC, *motive* tends to display both positive and neutral semantic prosody.

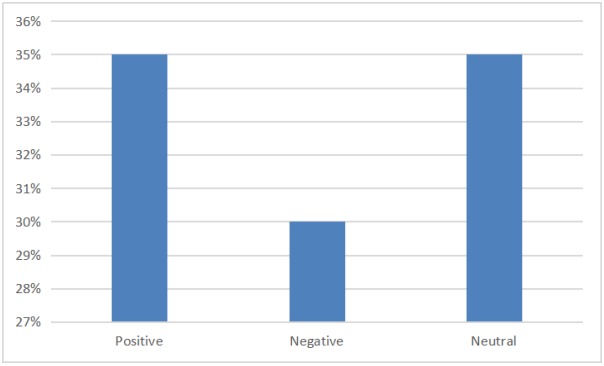

Figure 3: The Semantic Prosodic Distribution of the Collocates for motive in CLEC

Concordances are extracted and shown in the following.

1. Without self-discipline, there is no **motive** of **study** to speak of.
2. Stop indulging men's **unkind motives** towards to the animals before the nature shows angry.
3. There is no doubt our **motives** are **good**.
4. Mercy **killing** is conducted from different **motives** for different purposes.
5. I think it all depends. I think the **motive** to become **famous** is very important.
6. They graduated from high school. The **motive** varies from individual to individual.
7. **Competition** is one of the **motives** to the development of society.
8. A culprit may have many **motives** to **violate** the law, such as money, power.
9. I had **ulterior motives** to write to the vice-head teacher.
10. The new condition will bring them **interest** and **motive** for the work.

Among 10 concordances above, line 2, 4, 8, 9 present negative semantic atmosphere while 1, 6, 7, 10 manifest positive semantic environment and the rest of concordance 3, 5, suggest neutral semantic environments.

4.3.1.4 Semantic prosody of *motivation* in CLEC

All frequently used collocations of *motivation* in Chinese EFL learner corpus are investigated and the 30 significant collocates are divided into three categories as follows.

Positive (18)	*competition, study, education, development, improvement, improve, compete, important, support, mercy, knowledge, interesting, university, work*
Negative (9)	*suicide, criminal, commit, killing, discrimination, lack, pressure, stress*
Neutral (3)	*Different, society, students.*

Table 9: The semantic prosodic distribution of the collocates for *motivation* in CLEC

Motivation			
Frequency in Total			30
Semantic prosody	Positive	Negative	Neutral
Frequency counts	18	9	3
Percentage	60%	30%	10%

It is noticed that *motivation* in CLEC gains positive semantic prosody by collocating with larger share of positive words which reach up to 60% of all significant collocates, while 10% of collocates are endowed with neutral meaning and 30% of them signal negative meaning.

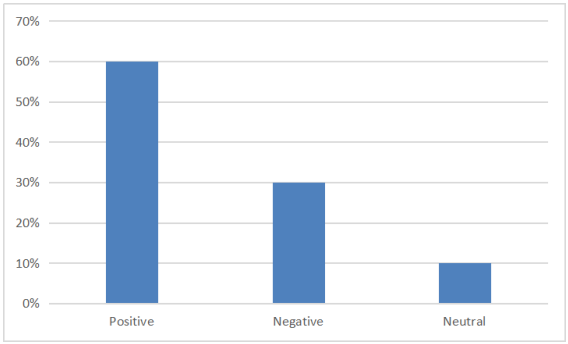

Figure 1 Figure 4: The Semantic Prosodic Distribution of the Collocates for motivation in CLEC

Some concordances are extracted to specify contextual information.

1. No competing, children will lose their **motivation** to **study**.
2. Without stress we may find no **motivation** in **work.**
3. Knowledge is the most important **motivation** for **development.**
4. The knowledge is the **motivation** of society **improvements.**
5. Students should learn to compete because **competition** is a **motivation** to urge them to

study.

6. Functions listed are the most people's **motivation** for going to the **university.**
7. Three parts: **criminal motivations**, the criminal's behavior and the harm of this.
8. End his life in a painless way has no **motivation** to **kill** a person who will die.
9. Need people to escape from their **evil motivation** or insatiable acquisitiveness.
10. Gaining **education** is a **motivation** to support one to survive.

Among the 10 concordances above, line 7, 8, 9 indicate negative semantic atmosphere whereas the rest of concordances reveal positive semantic atmosphere and concordances that involve with neutral semantic environment are not found.

4.3.1.5 A contrastive analysis of semantic prosody of *motive* and *motivation* in COCA and CLEC

Table 10: Semantic prosody of *motive* and *motivation* in CLEC and COCA

Synonyms	Corpora	Positive	Negative	Neutral	Semantic Prosody
motive	CLEC	35%	30%	35%	Mixed
	COCA	25%	52%	23%	Negative
motivation	CLEC	60%	30%	10%	Positive
	COCA	49%	6%	45%	Mixed

As observed in the above Figure, *motive* exhibits negative semantic prosody in COCA whereas *motive* in CLEC tends to collocate with positive, negative and neutral lexical items. *Motivation* is apt to collocate with words denoting positive and neutral affective meanings in COCA whereas *Motivation* displays positive semantic prosody in CLEC.

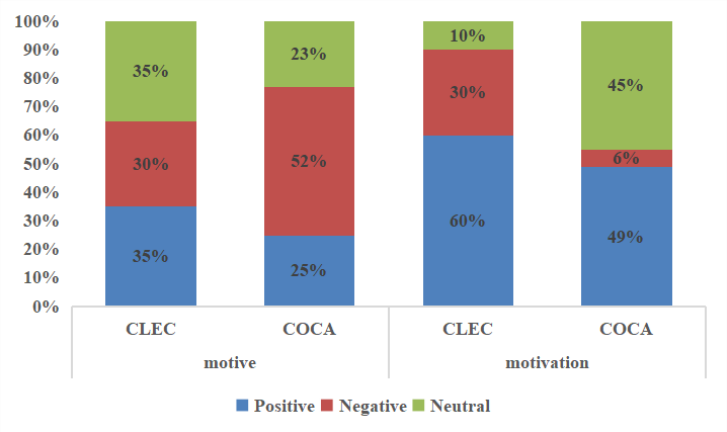

Figure 5: Semantic prosody of motive and motivation in CLEC and COCA

Thus, it is not difficult to see that semantic prosody of *motive* and *motivation* employed by Chinese EFL learners are to a great extent, misused. For the word *motive,* in line with what has been discussed previously, Chinese EFL learners are apt to underuse negative semantic prosody of *motive* for 20% more collocates with negative attitudinal meanings are not detected in CLEC(30%), contrasted to COCA(52%). Additionally, Chinese EFL learners often times, overuse the positive and neutral collocates of *motive.* Regarding the word *motivation*, we can find that Chinese EFL learners tend to overuse its negative semantic prosody and underuse its neutral semantic prosody for only 10% significant collocates in CLEC deliver neutral attitudinal meaning compared with 45% in COCA. It also could be detected that the significant collocates showing negative affective emotions reaching up to 30% in CLEC whereas only 6% could be observed in COCA.

Those collocational differences demonstrate that Chinese EFL learners are, more often than not, violate the rules of semantic prosody, thus leading to communicative failures via employing atypical collocates. To be more specific, in CLEC, interlanguage collocations and unusual collocations (Wei, 2006) such as *intrinsic motive, criminal motivation*, *motive to develop* are inconsistent with typical collocates preferred by native speakers. Some atypical or unusual collocations from concordance lines are shown as follows.

Table 11: Atypical collocations of *motive* and *motivation*

1 .We can say that competition is one of the **motives** of the **development** of society. When it comes to ... CLEC.txt 03-126
2. our awareness of competing is, the stronger our **motive** to make **success** is. The wheels of history don't ... CLEC.txt 01-18

3. Illegal killings which led by the **criminal**'s own **motivation** and disregard the patient's willing should be... CLEC.txt 146-32
4. Not long ago, after we know her **motivation** that she will **commit** suicide, the friends become. CLEC.txt 117-89

These 4 lines, apparently, breach the rules of semantic prosody and cross the border of semantic sets. The first 2 examples relate *motive* with *development* that demonstrates strong positive semantic preference. Likewise, example 3 and 4 associate *motivation* with "criminal, illegal killing or commit suicide", suggesting the confusion or misuse of these two near synonyms by Chinese EFL learners.

5. Conclusions

This corpus-based contrastive study investigates the use of two pairs of synonyms "motive" and "motivation" in terms of their frequency, collocation features and semantic prosody between Chinese EFL learners and native English speakers. The findings are summarized as follows.

Firstly, frequency counts of near synonyms in two corpora showcase that both English native speakers and Chinese EFL learners are inclined to employ *motivation* compared with *motive* in their choices of expressions.

Secondly, significant collocates of *motivation* tend to relate with "beneficial elements" for participating in rewarding activities while significant collocates of *motive* have a tendency to go with "evil intentions, crimes and irritating behaviors" in COCA. However, negative sense of "criminal and killing" that falls into semantic sets of *motive* in COCA are also detected as shared significant collocates of *motive* and *motivation* in CLEC, indicating that Chinese EFL learners could not distinguish the two words well when it comes to the lexical choice for expression.

Thirdly, it is identified that even synonyms enjoy similar meanings and share identical collocates, their semantic prosody unfolds different features. For example, in COCA, *motive* bears negative semantic prosody whereas *motivation* displays mixed semantic prosody. In CLEC, *motive* reveals mixed semantic prosody while *motivation* demonstrates positive semantic prosody. This indicates that Chinese EFL learners are prone to underuse negative semantic prosody of *motive,* but overuse its positive and neutral semantic prosody; in the meantime, they are apt to overuse positive and negative semantic prosody of *motivation* but underuse the its neutral semantic prosody to a certain degree.

In line with the idea that "Without grammar very little can be conveyed; without vocabulary nothing can be conveyed" (Wilkins, 1972, p. 111), vocabulary has been always been prioritized in language learning in China. Nevertheless, language transfer, lack of pragmatic awareness and old-fashioned methods for teaching and learning vocabulary may to a large extent, account for the misuses, overuses or underuses of English synonyms. In fact, Chinese EFL learners, when looking for a word for expression, are prone to employ word-for-word translation with the help of Chinese equivalents, thus treating synonyms as identical and replaceable with each other. Furthermore, Chinese EFL learners fail to develop the pragmatic awareness in communication due to insufficient exposure to natural communicative environment for using English. In addition, learners are always taught mainly on linguistic forms of vocabularies like lexical meaning and grammatical usage, and in the meanwhile, lack of sufficient input of semantic prosodies of a lexical item.

In view of the aforementioned, we claim that the approach of corpus-based analysis of synonyms should be recommended to learners so as to cultivate their habit of using corpus as an assistant rather than rely on their prior collocational experiences when confused with synonyms. What is more, language teachers are supposed to integrate semantic prosody into vocabulary teaching to help them avoid pragmatic errors and enhance communicative competence. In addition, EFL textbook glossaries need to provide semantic prosodies for lexical items. Finally, it would be beneficial to incorporate semantic prosody into lexicography with the purpose of illustrating idiomatic and typical collocational behaviors of synonyms.

6. Acknowledgments

This study is funded by Teaching Reform Project of NWPU 2018 and Higher Education Research Fund of NWPU (G2018KY0207).

References

Anthony, L. AntConc. Downloaded from http://www.laurenceanthony.net/

Channell, J. (2000). Corpus-based analysis of evaluative lexis. In S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 39-55). Oxford: Oxford University Press.

Firth, J. (1957). *Papers in linguistics, 1934-1951.* London: Oxford University Press.

Gries, S. T. (2008). Phraseology and linguistic theory: A brief survey. In S. Granger & F. Meunier (Eds.), *Phraseology: An Interdisciplinary Perspective* (pp.3-25) Amsterdam: John Benjamins

He, X.D (2003). *Dictionary of English Synonyms.* Beijing: The Commercial Press.

Hoey, M. (1991). *Patterns of lexis in text.* Oxford: Oxford University Press.

Hoey, M. (2000). A world beyond collocation: New perspectives on vocabulary teaching. In M. Lewis (Ed.), *Teaching collocations: Further developments in the lexical approac* (pp. 224-245). Hove, UK: Language Teaching.

Huang, R.H., (2007). A corpus-based study of semantic prosody of adjectives amplifiers of Chinese English learners. *Foreign Language Education, 28 (5), 57-60.*

Hunston, S. &Thompson, G. (2000). *in Text: Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press, 2000.

Hunston, S. (2002). *Corpora in applied linguistics.* Cambridge: Cambridge University Press.

Kjellmer, G. (1987). Aspects of English Collocations. In W. Meijs (Ed.), *Corpus Linguistics and Beyond. Proceedings of the seventh international conference on English language research on computerised corpora* (pp. 133-140.). Amsterdam: Rodopi.

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.

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: University of Birmingham.

Lu, J. (2010). A corpus-based study on collocational behavior and semantic prosody in Chinese English learners. *Modern Foreign Languages, 33*(3), 276-285.

Partington, A. (1998). *Patterns and meanings: Using corpora for English language research and teaching.* Philadelphia: John Benjamins.

Partington, A. (2004). Utterly content in each other's company: Semantic prosody and semantic preference. *International Journal of Corpus Linguistics, 9*(1), 131-156

Sardinha, T. (2000). Semantic prosodies in English and Portuguese: A contrastive study. *Cuadernos de Filologia 9*(1), 93-110.

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Sinclair, J. (1996). The search for units of meaning. *Textus: English Studies in Italy,* 9, 75-106.

Sinclair, J. (2000). Lexical grammar. *Naujoji Metodologija, 24,* 191-203.

Sinclair, J. (2004). *Trust the Text: Corpus and Discourse.* London: Routledge,

Stubbs, M. (1995). Collocations and semantic profiles: On the cause of trouble with quantitative studies. *Functions of Language,* 2(1), 23-55

Stubbs, M. (1996). *Text and CorpusAnalysis.* Oxford: Blackwell

Stubbs, M. (2009). The search for the unit of meaning: Sinclair on empirical semantics. *Applied linguistics, 1*, 115-137.

Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics.* New York: Blackwell.

Sun, L. (2017). A new approach to investigate the features of the synonyms' semantic prosody of the Chinese EFL learners: A contrast between "Affect" and "Influence". *Modern Foreign Language Studies,1*, 57-61.

Tognini-Bonelli, E.(2001). *Corpus Linguistics at Work*. Amsterdam: John Benjaxnins.

Wilkins, D. (1972). *Linguistics in Language Teaching.* London: Hodder & Stoughton Educatinnal.

Wei, N.X. (2002). A corpus-driven study of semantic prosodies in specialized texts. *Modern Foreign Language,* 25(2), 165-175

Wei, N.X. (2006). A corpus-based contrastive study of semantic prosodies in learner English. *Foreign*

Language Research, 132, 50-54.

Xiao, Z., & McEnery, A. (2006). Collocation, semantic prosody and near synonymy: A cross-linguistic perspective. *Applied Linguistics, 27*(1), 103-129

Zethsen, K.K. (2006). Semantic prosody: Creating awareness about a versatile tool. *Tidsskrift for Sprogforskning, 4*(1), 275-294