EFL Learners' Synonymous Errors: A Case Study of *Glad* and *Happy*

Yingying Shen
Department of Foreign Languages
Dongguan Nanbo Polytechnic
Dong Guan 523083, China
E-mail: selena121212@yahoo.com.cn

Abstract—Corpus linguistics and its relationship with modern lexicography are paid more and more attention by linguists and lexicographers in recent years. This paper, based on corpus analysis, mainly studies EFL (English as Foreign Language) learners' synonymous errors (SE), which refer to errors made by learners when using synonyms in writing. The unit observed is a pair of commonly used synonyms glad and happy, and the discussion includes both the phenomenon and reasons. Moreover, this paper considers that constructing a semantic network of synonyms in LDs (learner's dictionaries) can simulate learner's mental lexicon and help them to prevent from SE. Therefore, this paper takes LDs (both electronic and printed ones) as examples to discuss how existing LDs differentiate synonyms by means of semantic network and whether there are some inadequacies.

Index Terms—synonymous errors (SE), corpus, semantic network

I. INTRODUCTION

In the course of the last couple of decades we have seen a breakthrough in the use of computer corpora in linguistic research, and they are used for a wide range of studies in grammar, lexis, discourse analysis, language variation, etc. (Granger, 2007:33) Corpus linguistics has played a crucial role in lexicography studies and provides valuable data for dictionary making and research.

English has many words that are considered synonymous, and Chinese EFL (English as Foreign Language) learners tend to mix them up and make certain kinds of errors. Thus, analysis of synonymous errors (SE) becomes a common concern for Chinese learners, teachers and L2 lexicographers in recent years. The objects observed in this paper are a pair of adjectives *glad* and *happy*, for they are commonly and frequently used in spoken or written English and seem almost identical in meaning for Chinese EFL learners.

Solutions to SE can be found in presenting synonyms by means of semantic network. Many argue that the organization of the mental lexicon is a network in which words/concepts are presented as nodes and are connected via relationships to other words in the network. (Carroll, 2000:109-110) Zhang Yihua presents a figure of "semantic network in mediostructure" in his latest work and states that how to combine semantic network (in mental lexicon) with the structure of dictionaries is the major task for active dictionaries. To construct a multidimensional semantic network in dictionaries can effectively activate learners' cognitive association and prior knowledge system. Therefore, the projection between new and old knowledge is realized in order to help learners understand and memorize new knowledge. (cf. Zhang Yihua, 2007) That is to say, a comparatively successful semantic network system of synonyms in LDs can not only simulate learners' mental lexicon but also provide guidelines for arranging synonymous information in dictionaries.

II. ERROR ANALYSIS

A. Learner Corpus

Learner corpus is an electronic collection of texts produced by foreign/second language learners, and we can find many interesting patterns of usage which are quite different from those of native speaker. In many cases, 'these differences are due to the fact that learners are still in the process of acquiring a language, and they naturally make errors and mistakes." (Tono, 2005:317)

Interest in learner corpus is growing fast and a number of learner corpora are being collected and analyzed nationally and internationally both at home and abroad. Chinese Leaner English Corpus (CLEC) has been designed and constructed by Guangdong University of Foreign Studies, Shanghai Jiao Tong University and other universities under the leadership of Professor Gui Shichun and Profeesor Yang Huizhong. Figure 1 shows its online search engine (English version):



Figure 1. CLEC online search engine

CLEC is a project of one million word corpus of English learners collected from five groups in China: high school students, CET-4 and CET-6 examinees, English majors at lower grades and English majors at higher grades, and is the major research corpus used in this paper. The corpus used for comparison is the Freiburg-LOB Corpus of British English (FLOB)¹, and tools used for the present study is Wordsmith Tools².

Glad and happy are often used to express someone's positive emotional mood and listed as synonyms in dictionaries. "Corpus-based lexicographic analysis are particularly well suited to uncovering systematic differences in the patterns of use", (Biber, 2000:43) and the present research is to study on the frequency and collocation of glad and happy used by Chinese EFL learners and then discuss the phenomenon, reasons, and solutions.

B. Overuse of Glad and Happy

The frequency of using *glad* and *happy* is significantly different between Chinese EFL learners and native speakers when these two words are observed in CLEC and FLOB. Table 1 shows that the frequency of *glad* (169) used by Chinese EFL learners (in CLEC) is over three times of that by native speakers (in FLOB), and the frequency of *happy* (812) is more than five times.

TABLE 1. FREQUENCY DISTRIBUTION OF GLAD AND HAPPY

Word	Freq (CLEC)	T×T%	Freq (FLOB)	T×T%	Keyness	
Glad	169	0.01	50		71.2	
Нарру	812	0.07	154	0.01	507.9	

As shown above, the *Keyness*³ of *glad* and *happy* (71.2 and 507.9) means these two words occur much more often than expected in comparison with FLOB. That is to say, Chinese EFL learners tend to overuse *glad* and *happy*.

C. Collocates of Glad and Happy

Collocations of these two synonyms are observed in CLEC and the Wordsmith tools can provide two kinds of information: keywords in context lines (KWIC) and frequent collocates of the keyword.

In CLEC, *Glad* is used 169 times, among which 125 are used as "be glad to...": variations of "be" are common left collocates of *glad*, which means "be glad..." is frequently used; some verbs such as "accept, join, tell, see" in R2 (the second one in the right side) shows that they seem to be used as infinitive form like "glad to accept, glad to join, glad to tell, glad to see etc." Usual modifiers of *glad* are "so" and "very".

The adverb "very", which occurs 231 times in L1 (the first one in the left side) of *happy*, is the most common modifier used, and Chinese EFL learners tend to use patterns such as "be/feel/become happy...". In R1, "for" appears 13 times, "that" 11 times "when" 10 times, and "with" 6 times, which illustrates that "be happy for/that/when/with..." is familiar to them. In addition, *happy* is used as a complement to nouns such as in "make her/me happy", as an attributive in "happy life, day, atmosphere...etc", as an intensifier in exclamatory sentence "how happy...!" or as a guide word in

1

¹ FLOB Corpus is one of similarly structured corpora spawned by the Brown Corpus in 1990s and used in this paper in order to figure out how native speakers use words in real context.

² Wordsmith Tool is an integrated suite of programs for looking at how words behave in texts and is used in this paper to compare words (e.g. *glad* and *happy*) used in CLEC and FLOB.

³ The positively *Keyness* means a word occurs more often than would be expected in comparison with the reference corpus, and a word with negatively *Keyness* occurs less often than would be expected.

⁴ Variations of "be" includes "was, are, were"

blessings like "Happy birthday".

Compared with Chinese EFL learners, *glad* is not so frequently used by native speakers (in FLOB) with its occurrence of 50. In FLOB, *glad* usually follows "was" and "I'm" as in sentences such as "be glad..."All of the 50 KWIC shows that *glad* is usually used by native speakers in patterns such as: "be glad to...", "be glad (that)...".

In FLOB, *happy* (freq. 154) is used by native speakers such as: "be happy with...", "be happy to...". Modifiers used to emphasize *happy* are "very, so, too, quite, unreasonably, particularly, extremely, blissfully, enough etc." In addition, *happy* is used as an attributive in collocations such as: "happy family, happy lady, happy period, happy one, happy marriage, happy endings etc."

The result of the corpus analysis can be shown in table 2, which gives a piece of information about SE made by Chinese EFL learners.

TABLE 2.
RESULTS OF THE STUDY ON USING *GLAD* AND *HAPPY*

Glad	RESULTS OF THE STUDIT ON USING GLAD AND HAPPT
	A + + 15 - 170
In CLEC	• total freq. 169
CLEC	 sentence patterns: be glad to (freq. 125); be glad (that) (freq. 14) not used before noun
	• modifiers: very, so, always, quite
	modifiers. very, so, aiways, quite
In	• total freq. 50
FLOB	• sentence patterns: be glad to (freq. 20); be glad that (freq. 22)
	 used as a complement: make someone glad
	 used before noun: glad rags
	 modifiers: so, very, clearly, just
Errors	1. overuse the word <i>glad</i>
	2. overuse "be glad to"
	3. ignore the use of <i>glad</i> as a complement in "make someone glad"
••	4. ignore the use of <i>glad</i> as an attributive
Happy	
In	• total freq. 812
CLEC	• sentence patterns: happy to, happy for, happy that, happy when, happy with
	 used as a complement: make someone/ something happy; give/ enjoy someone happy
	 used before noun: happy life, happy festival, happy night, happy time, happy day, happy family/ families,
	happy atmosphere etc.
	 modifiers: very, so, some, too, quite, be full of, more, most, great etc.
In	• total freq. 154
FLOB	• sentence patterns: be happy with, be happy to
	 used as a complement: make someone happy
	 used before noun: happy family, happy lady, happy period, happy marriage, happy endings etc.
	 modifiers: very, so, too, quite, unreasonably, particularly, extremely, blissfully etc.
Errors	1. overuse the word <i>happy</i>
	2. misuse "make something happy" "give/ enjoy someone happy"
	3. misuse collocations such as: happy festival, happy gathering, happy generation, happy old home etc.
	4. ignore the use of some modifiers such as "unreasonably, particularly, extremely, blissfully, etc."
	Note: the patterns and collocations in point 2 and 3 are searched again in BNC (British National Corpus) and the
	result is "empty", which means that native speakers of English seldom use such kinds of collocations.

D. Reasons of Making SE

The phenomenon of SE is studied in the preceding section on the basis of corpus comparison between CLEC and FLOB, while reasons of these SE are further discussed, which can be divided into two main categories: reasons on the overuse and reasons on errors in collocates.

1. On overuse of glad and happy

According to Wang, if the total English vocabulary can be estimated as 500,000 words, 2000 of them are core vocabularies and about 3000 to 5000 are basic vocabularies. (Wang Rongpei, 2000) He has divided general language vocabulary into several types: core vocabulary, primitive vocabulary, basic vocabulary and general vocabulary, all of which can be better illustrated in figure 2.

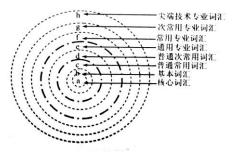


Figure 2. Classification and distribution of vocabulary (Zhang Yihua, 2007)

Core and basic vocabularies, which are also considered as high frequency words, are in the heart of the whole vocabulary system. High frequency words can be learned and mastered before lower frequency words from the cognitive perspective. Therefore, *Longman Dictionary of Contemporary English* uses 2000 basic and core vocabulary to define all the headwords, for the writers believe that these basic vocabularies and high frequency words can be easily learned and understood by English learners.

In BNC frequency list, *glad* occurs in No. 2700 and *happy* in No. 924, which means that both of them are within the 3000 basic vocabularies. Therefore, it is easier for Chinese EFL learners to learn and master basic words *glad* and *happy* than any other special or uncommon words that seem more difficult for them. When they are about to express their pleasant feelings, they tend to use *glad* and *happy* much more often. That's the reason why *glad* and *happy* are overused by Chinese learners when compared with native speakers.

2. On collocates of glad and happy

The transfer of native culture is first put forward by Rod Ellis in 1997 and attracts more and more linguists' interests in recent years. Wang concludes three kinds of transfer in the process of second language acquisition (SLA): positive transfer, negative transfer and zero transfer. (Wang Yongde, 2008) When grammatical rules or vocabularies of second language are the same as or similar with those of native language, positive transfer occurs and learner's SLA is facilitated. Contrarily, negative transfer occurs when second language differs significantly from learner's native language and their native prior knowledge in the mental lexicon cannot match perfectly with the second language. In that case, learners are likely to either misuse some patterns and collocations when they fail to tell the difference between second language and native language or overuse some grammatical rules when they consider those special rules of second language as general ones. Thus, errors such as misuse and overuse some patters, collocations or modifiers of glad and happy are made by Chinese EFL learners. But, if the grammatical rules of second language are too difficult for them to acquire, they tend to shy away from using some words, patterns or collocations in communication, and this phenomenon is so-called zero transfer. For example, in communication or writing, they may use common and simple words glad and happy instead of specific and difficult words such as cheerful, in a good mood, light-hearted, delighted, and gloating, etc. Likewise, they may use familiar patterns and collocations to produce sentences or discourses, and prevent themselves from using patterns or collocations they do not know. This will lead to underuse of some patterns, collocates and modifiers that are frequently used by native speakers. Table 3 concludes the influence of native language transfer on SE of glad and happy.

TABLE 3.

NATIVE LANGUAGE TRANSFER AND ERRORS OF GLAD AND HAPPY

Transfer of native culture	Results	Examples	
positive	Facilitation of learner's SLA		
negative	Misuse or overuse of words, patterns, collocations, etc.	Overuse "be glad to" Misuse "make something happy" Misuse collocations such as "happy gathering, happy generation, happy old home etc."	
zero	Underuse of words, patterns, collocations, etc.	Underuse synonyms such as cheerful, light-hearted, delighted etc. Ignore the use of "make someone glad"	

III. SOLUTION TO SE: SEMANTIC NETWORK OF SYNONYMS IN LDS

The development of corpus linguistics enables linguists and lexicographers to research words from a new perspective and "the consequence and requirement of combining dictionaries with corpus study is to develop lexicography". (Zhang Yihua, 2004:43) Therefore, the study of SE based on corpus research mentioned in preceding sections is used for more comprehensive discussions about lexicography. Since the target users of LDs (learner's dictionaries) are language learners, how to solve problems of users' language learning and how to help them in real sense is a common concern for

both researchers and lexicographers. The author considers that a comparatively better way to help Chinese EFL learners' synonym acquisition is to build a semantic network of synonyms in learner's dictionaries.

Synonymous relationship between words forms the basis of traditional definition and semantic network in mental lexicon. (Zhang Yihua, 2002: 179) In the light of mental lexicon, to construct a semantic network in LDs can: provide similarities and differences between synonyms for dictionary users, enlarge EFL learners' vocabulary acquisition of synonyms, and activate learners' cognitive concepts and semantic association in mental lexicon.

A. Semantic Network of Synonyms in Electronic LDs

In this section, semantic network of *glad* and *happy* are probed into in electronic LDs: Longman Dictionary of Contemporary English (CD-ROM 4th edition, LDCE for short) and Oxford Advanced Learner's Dictionary (CD-ROM 7th edition, OALD for short).

1. In LDCE:

Glad is defined as "[not before noun] pleased and happy about something" and its collocations and examples are listed below the headword. The "Longman Language Activator" facilitates users to activate the node in their mental lexicon and finding out the words closely related to the headword (node) through semantic network. Figure 3 shows the "Activate your language" window which firstly links to "to be willing to do something" then to "be glad/ happy/ pleased to do something".

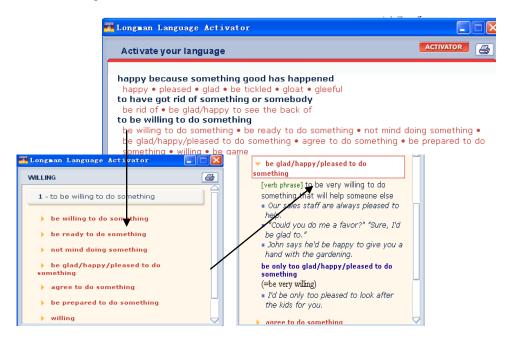


Figure 3. "Activate your language" under the headword glad

In addition, users can click "happy because something good has happened" in the above "Activate your language" window and reach to a usage note under the headword *happy*, which realizes the co-relation between the entries *glad*, *happy*, and other related synonyms.

Similarly, *happy* is searched in LDCE, and the pop-up window offers a number of patterns of use which can link to various synonyms of *happy*: "cheerful, contented, jolly, blissful, idyllic...etc", and these patterns are listed under different headwords. Thus, a semantic network that connected a number of entries is constructed in the dictionary body.

2. In OALD

Under the headword *glad*, definition, examples, idioms, structures and word origin of *glad* are provided and a pop-up window on synonyms comes up as in figure 4. In this window, illustrations of a group of synonyms are presented. Under the guide words "glad, happy or pleased?" error warnings of these three synonyms are briefly given. Besides, the "Word finder" function enables users to reach another "word finder" under the entry of *happy*, which explains the usage of both *happy* and its synonyms such as "delighted" "cheerful" "pleased"... etc. Double clicking any of these synonyms can bring users to another entry under the headword "delighted" "cheerful" and so on.

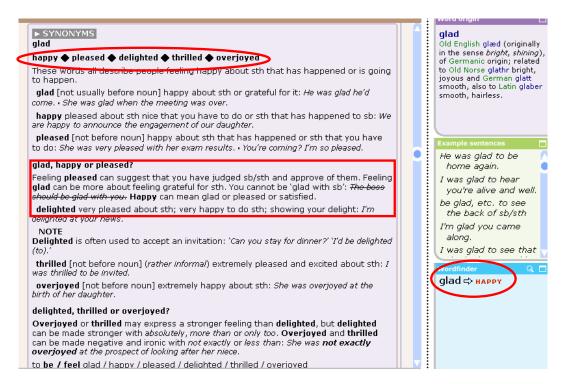


Figure 4. Synonyms and Word finder under the headword glad

3. Comments:

Electronic dictionaries such as LDCE and OALD have constructed a comparatively mature semantic network of synonyms by means of "Longman Language Activator" or "Word finder", which can facilitate users' language acquisition.

However, the error warnings or discrimination of synonyms are not particularly designed for Chinese learners. In order to meets the needs of Chinese EFL learners, it is necessary for lexicographers in China to consider how to make use of useful information gained from corpus studies. What's more, certain exercises about sentence patterns, modifiers or collocations can be added in electronic dictionaries in order to help Chinese EFL learners in SLA.

B. Semantic Network of Synonyms in Printed LDs

Semantic network of synonyms is also practiced by printed LDs, and dictionaries investigated are: Cambridge Advanced Learner's Dictionary (CALD), Collins COBUILD Advanced Learner's English Dictionary (COBUILD), MacMillan English-Chinese Dictionary (MECD), A New English-Chinese Dictionary ("New Century"), New Age English-Chinese Dictionary ("New Age") and A Modern English-Chinese/Chinese-English Dictionary ("Modern Dictionary").

In CALD, COBUILD and "New Century", no cross-reference or synonymous discrimination is included in the entries of *glad* and *happy* while in MECD, under the headword *happy* a special column called "words you can use instead of happy" is presented and any synonym shown in the column can direct users to look it up in the same dictionary and find detailed information. The classifications of *happy* used in different situations can guide users to learn more about various synonyms of *happy* and choose a proper word to express a specific meaning in particular context. Moreover, the relation between *happy* and its synonym such as "cheerful" or "light-hearted" can form a semantic network system in MECD.

Likewise, in "New Age" and "Modern Dictionary", special column and usage note are given as in figure 5 and figure 6, which illustrates synonyms of *happy* and *glad* respectively.

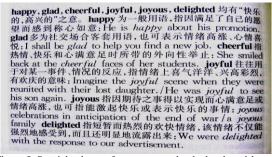


Figure 5. Special column of synonyms under the headword happy



Figure 6. Usage note under the head word of glad

Printed dictionaries, although have the limitation of dictionary space, has put some effort to semantic network of synonyms building. However, there are still some dictionaries such as CALD, COBUILD and "New Century" that pay no attention to synonymous network or discrimination.

The ways in which they provide synonyms are limited, either by special columns or usage notes. The results of corpus research and dictionary analysis mentioned can lead us to conclude that it is necessary for Chinese lexicographers to consider whether there are other ways to present synonymous information, whether error warnings about synonyms should be included in printed LDs, and how to improve the semantic representation of synonyms in printed LDs.

IV. CONCLUSION

Corpus linguistics is closely related to modern lexicography. The main purpose of this paper is to study Chinese EFL learners' SE on the basis of corpus comparison in order to provide some implications to dictionary making. Findings of the present study can be concluded as follows.

First, based on comparison between CLEC and FLOB, SE of *glad* and *happy* made by Chinese EFL learners can be concluded as: overuse of *glad* and *happy*, overuse or underuse of some patterns, and misuse of some modifiers or collocations.

Second, overuse of *glad* and *happy* occurs because both of them belong to Basic English Vocabularies, which are easier to be learnt and used. Misuse, overuse and underuse some patterns, collocates and modifiers of *glad* and *happy* result from "transfer of native language".

Third, Chinese EFL learners tend to make some errors in the process of SLA, so LDs have to present sufficient synonymous information for their users in the form of semantic network. Both electronic and printed dictionaries, more or less, attempt to present synonyms by means of semantic network. However, Chinese lexicographers need to further consider questions such as: whether semantic network is the most effective way to help learners to acquire synonyms? To what extent the semantic network of synonyms helps them? How to modify the semantic network in learner's dictionaries? And so on.

REFERENCES

- [1] Longman Dictionary of Contemporary English UPDATED EDITION (CD-ROM 4th edition) Longman Group UK Ltd 2004
- [2] Oxford Advanced Learner's Dictionary (CD-ROM 7th edition) Oxford University Press 2005
- [3] Cambridge Advanced Learner's Dictionary (2nd edition) CAMBRIDGE University Press 2005
- [4] Collins COBUILD Advanced Learner's Dictionary (5th edition) Foreign Language Teaching and Research Press 2006
- [5] MacMillan English-Chinese Dictionary for Advanced Learners, Foreign Language Teaching and Research Press 2005
- [6] A New English-Chinese Dictionary (Century edition), Shanghai Translation Publishing House 2000
- [7] New Age English-Chinese Dictionary, The Commercial Press 2004
- [8] A Modern English-Chinese/Chinese-English Dictionary, Foreign Language Teaching and Research Press 2005
- [9] Granger, S. (2007). Corpus-Based Approaches to Contrastive Linguistics and Translation Studies. Beijing: Foreign Language Teaching and Research Press, 33.
- [10] David W Carroll. (2000). Psychology of Language (3rd edition). Thomson, 109-110.
- [11] Tono, Y. (2005). The Potential of Learner Corpora for Pedagogical Lexicography. Words in Asian Cultural Contexts: 315-321.
- [12] Biber, D. (2000). Corpus Linguistics. Beijing: Foreign Language Teaching and Research Press & Cambridge: Cambridge University Press, 43.
- [13] Zhang, Yihua. (2009). Yuyi Renzhi Shiyi (Semantics Cognition Definition). Shanghai: Shanghai Foreign Language Education Press, 180-182.
- [14] Wang Rongpei. (2000). Standard English and Core Vocabulary. Foreign Language Teaching and Research, (12): 5
- [15] Zhang, Yihua. (2007). Dangdai Cidianxue (Contemporary Lexicography). Beijing: Commercial Press, 59.
- [16] Wang Yongde (2003). On the Role of Mother Tongue in Second Language Acquisition. *Journal of Wuyi University (Social Sciences)* 5 (3).
- [17] Zhang Yihua. (2004). Computational Lexicography and New-Typed Dictionaries. Shanghai: Shanghai Lexicographical Publishing House, 43
- [18] Zhang, Yihua. (2002). Yuyixue yu Cidian Shiyi (Semantics and Definition in Dictionaries). Shanghai: Shanghai Lexicographical Publishing House, 179.