

A Corpus-based Study of Chinese English Learners' Use of *Somewhat**

Qian Yang

Shandong University of Finance and Economics, Ji Nan, Shandong Province, China

Abstract—This study investigates Chinese English Learners' usage of one of the downtoners---*somewhat* in an attempt to explore the features and development pattern of Chinese English Learners' usage of this downtoners. The study reveals that the Chinese English Learners differ greatly from the native English speakers in frequency, collocation, colligation, semantic preference and semantic prosody. In order to improve the Chinese English learners' language proficiency, it's imperative for the teachers to change the traditional English teaching approach and adopt the corpus-driven teaching method.

Index Terms—downtoners, written language, somewhat, extended meaning unit

I. INTRODUCTION

Intensifiers are the words that serve as modifiers of adjectives, adverbs, verbs and prepositional phrases to describe the state or degree of the topics being talked about and intensify the meaning expressed in these words. According to Quirk, there are two types of intensifiers, amplifiers and downtoners.

Amplifiers scale upwards. They are divided into (a) maximizers, which can denote the upper extreme of the scale, and (b) boosters, which denote a high degree, a high point on the scale. Downtoners have a generally lowering effect on the force of the verb or predication and many of them apply a scale to gradable verbs. They can be divided into four groups: (a) approximators serve to express an approximation to the force of the verb, while indicating that the verb concerned expresses more than is relevant. (b) compromisers have only a slight lowering effect and tend to call in question the appropriateness of the verb concerned. (c) diminishers scale downwards and roughly mean 'to a small extent'. (d) minimizers are negative maximizers, '(not) to any extent'. According to Quirk, *somewhat* can be classified as a diminisher.

As to the study of the units of meaning, most of the primitive studies regard word as the basic unit of language. Sinclair(2004) pointed out that there are many cases in texts where the independence of the choice of words is compromised, because other patterns cut across them and constrain them. That is, some words predict the appearance of other words because they are frequently co-occurrent. He further proposed the extended meaning unit, which combine collocation, colligation, semantic preference and semantic prosody. Collocation is a frequent co-occurrence of words; colligation is the co-occurrence of grammatical choices (Firth) to account for the greater variation; semantic preference refers to the meaning of the words frequent collocates; semantic prosody is attitudinal, and on the pragmatic side of the semantics/pragmatics continuum. Semantic prosody has a leading role to play in the integration of an item with its surroundings. It expresses something close to the 'function' of the item. The extended meaning unit changes the traditional linguists' practice of isolating words from grammar; it removes the research focus from the actual words to the grammatical features co-occurrent with the word; it is a more complex and comprehensive unit combining semantics with pragmatics.

II. LITERATURE REVIEW

With the development of the corpus linguistics, intensifiers attract the attention of many corpus linguists' attention. Granger (1998) made an investigation on the collocation of the intensifiers ended with -ly. Lorenz (1998) studied the adjective intensification of the Germany English learners. Kennedy (2003) found that different maximizer will co-occur with different words.

Greatly influenced by these studies, Chinese linguists began to make contrastive investigation on the use of intensifiers of Chinese English learners. Maocheng Liang (2003) analyzed the intensifiers in the oral production of Chinese English learners in a corpus of Chinese EFL learners' spoken English and compares it with that in an English native speakers' corpus. His corpus-based analysis reveals that Chinese EFL learners use the booster word very far too much, and this results in a tendency of overstatement. Zhang Xia (2005) investigates the intensifiers in the Chinese non-English major EFL learners' oral production in COLSEC. Her study reveals that Chinese non-English major EFL learners tend to overuse, underuse and misuse many intensifiers compared with the native speakers. Wang Haihua and Chen Guohua (2007) pointed out that high-school Chinese EFL learners overuse the intensifiers like so, very and much.

* This paper was supported by the foundation of Shandong University of Finance and Economics with the project No. jy 201255.

They also found that the variety and quantity of the intensifiers used by these learners will greatly increase as their language proficiency improves.

At present, studies on the intensifiers have the following two drawbacks: 1. majority of the previous studies centered on amplifiers while neglecting downtoners. 2. most linguists study the use of intensifiers in oral production while investigation on the intensifiers in written language are rare. The application of intensifiers is one of the important factors in EFL learners' language proficiency, thus studies on these intensifier play a significant role in the investigation of learners' interlanguage. Studies on downtoners, an indispensable part of the intensifiers, are far from enough considering from the depth and width of the previous studies. The present study will investigate the use of *somewhat*, one of the downtoners, in the Chinese EFL learners and compare it with the native speakers.

III. METHODOLOGY

A. Research Question

This study sought to find out answers to the following research question:

- a. Is there a significant difference in the frequency of *somewhat* in written language between the Chinese EFL learners and the native speakers?
- b. Is there a significant difference in the core collocation, colligation, semantic preference and prosody of *somewhat* between the Chinese EFL learners and the native speakers?
- c. Is there a significant difference in the use of *somewhat* among the Chinese EFL learners in different levels?

B. Instruments

The native English corpus used in the present study is the BYU-BNC developed by Davies Mark with the BNC (British National Corpus) developed by the Oxford University Press as the basis. The BYU-BNC has over 1,000,000 million words, and is composed of two parts: written language and oral language. The learners' English corpus in this study is the Chinese Learners' English Corpus (CLEC). CLEC, with over 1,000,000 words, is developed by Gui Shichun and Yang Huizhong. It is composed of five sub-corpus, each sub-corpus belongs to a specific level of the Chinese English Learners. (st2: middle school; st3: the students that passed College English Test Band 4; st4: the students that passed College English Test Band 6; st5: the freshmen and sophomore of the English majors; st6: junior and senior students of the English majors). In this study, the five sub-corpus are re-classified into 3 groups, that is, primary English learners—st2, medium English learners—st3 and st4; advanced English learners—st5 and st6. Antcon developed by Laurence Anothony of the Waseda University is used as the searching instrument of the present study.

C. Procedure

- Step 1: search the node—*somewhat*, and compare the frequencies of the word in the two corpora.
- Step 2: search the significant collocations of *somewhat* in the two corpora and make a contrastive analysis.
- Step 3: conclude the colligation of *somewhat* in the two corpora and make a contrastive analysis.
- Step 4: conclude the semantic preference of *somewhat* in the two corpora according to the significant collocations.
- Step 5: make a contrastive analysis of the semantic prosody of *somewhat* in the two corpora.
- Step 6: conclude the features and development patterns of the use of *somewhat* of Chinese English learners of different language proficiency.

D. Data Analysis

1. Frequencies

	BNC	CLEC	Chi-square	Significance (p)
raw frequency	4317	15		
standard frequency (per million)	50.0349	14.0108	27.6730	0.000***

As is shown in Table 1, Chinese English Learners use significantly less *somewhat* compared with the native English users, which coincides with the research result of the previous studies of Maocheng Liang (2003) and Xia Zhang (2005). The reasons of the underuse of *somewhat* may be: a. Most of compositions in the CLEC, the topics of which are given beforehand, are argumentative essays. In order to make the essays more argumentative, the students will try to use more maximizers to intensify the mood, consequently they will use less downtoners. b. *Somewhat*, according to the English-Chinese dictionary, means“有点, 稍微”, and there are many other words and phrases that have the similar meaning, like *a little*, *a bit*, *kind of*. *Somewhat* is the most difficult word in this group of synonyms, and it is acquired later than other synonyms according to the teaching syllabus of China. So when the student wants to express “有点、稍微”, they will try to avoid the use of *somewhat* for making less mistakes.

2. Significant Collocations

The first 50 significant collocations of *somewhat* in BNC are listed in Table 2. Among the 50 significant collocations, there are 24 adjectives, 18 verbs (with 13 past participles and five present participles), four adverbs and three nouns. Because past and present participle can be used as verb and also as adjective to denote the condition, the concordance

lines will be further analyzed in the present study.

TABLE 2:
SIGNIFICANT COLLOCATIONS OF *SOMEWHAT* IN BNC (FIRST 50)

order	collocation	frequency	MI score	order	collocation	frequency	MI score
1	different	180	3.42	26	remote	13	3.60
2	similar	76	3.45	27	dubious	12	5.52
3	surprisingly	32	5.02	28	obscure	12	4.73
4	larger	29	3.33	29	recovered	12	3.87
5	unusual	26	4.11	30	superficial	11	5.22
6	arbitrary	25	5.93	31	complicated	11	3.35
7	albeit	23	5.45	32	ironic	10	5.22
8	surprising	23	4.11	33	reluctant	10	3.74
9	vague	22	5.34	34	artificial	10	3.22
10	confused	22	4.65	35	variable	10	3.22
11	manner	22	3.27	36	blurred	9	5.33
12	surprised	21	3.55	37	contradictory	9	5.06
13	differently	20	5.20	38	bizarre	9	4.52
14	easier	19	3.27	39	crude	9	4.12
15	aback	18	7.25	40	disappointed	9	3.53
16	surprise	17	3.17	41	tended	9	3.13
17	ambiguous	15	5.56	42	isolated	9	3.08
18	uncertain	15	4.29	43	esoteric	8	6.59
19	misleading	14	4.82	44	bemused	8	6.13
20	restricted	14	3.55	45	simplistic	8	5.86
21	confusing	13	5.42	46	shaky	8	5.55
22	reluctantly	13	5.20	47	strained	8	4.81
23	disappointing	13	5.12	48	slower	8	4.55
24	lacking	13	4.52	49	neglected	8	4.15
25	modified	13	4.21	50	odds	8	3.89

Quirk(1985) pointed out that some adjectives look like participles with the suffix –ing or –ed can be classified as participial adjectives. These participial adjectives present static meaning, which are not connoted in their corresponding verbs. The standard of judging a participle is whether an adjective or a verb is to use the word *seem* to substitute the linking verb *be*, because only the *be* followed by adjectives can be substituted by *seem*.

In the present study, the researcher uses this standard to analyze all the concordance lines. The research shows: a. There are altogether 76 present participles in all the concordance lines, and all these participles are used to denote the state of something or an affair. b. There are 153 past participles in all the concordance lines. 41 of them are used to denote an action, and the remaining 112 items are used to signify the outcome or the state of something or an affair. According to the above-mentioned analysis, a conclusion that *somewhat* is usually used as downtoners to lower the meaning of the adjectives denoting the nature and state of a thing or an affair.

There are 15 concordance lines of *somewhat* in CLEC; they are listed in Table 3. Five of them are used to modify verbs, another five are used to modify adjectives, two for modifying preposition, two for preposition phrases and one for explaining the meaning of the whole sentence.

TABLE 3:
CONCORDANCE LINES OF *SOMEWHAT* IN CLEC

1	English word, sometimes, it is	somewhat	difficult for you to pronounce, time by t
2	spend more time on lessons. So	somewhat	we are more busy than before. In
3	for the job. Even if they are	somewhat	unsuitable to it, they convince that the
4	and light rarely patronized the	somewhat	Stygian staircase; everyone was apathetic
5	year old small black old suitcase,	somewhat	like school bag. At the beginning of the
6	middle of a park. Restaurant A is	somewhat	like a classical restaurant. Oil paintings
7	but patient as well. They seem	somewhat	hopeful but have to wait ilently outside
8	ets with money; however, it also	somewhat	ruined our lives. Actually, these damages
9	ich we easily neglect and which	somewhat	influence our studies. First, is interest.
10	views upon life and death have	somewhat	altered. Still, Chinese people believe that
11	nned, on the contrary, it produces	somewhat	criminals. What's wrong? To punish the c
12	rthermore, though a degree might	somewhat	appears theoretical, actually it values
13	certification, a certain degree is	somewhat	a kind of proof of a graduate's educational
14	they like them. Military service is	somewhat	compulsory. Some young people are
15	As to the people, their taxes can	somewhat	be lessened for the military budget goes

There is a striking difference in the collocates of *somewhat* between the Chinese English learners and the native English speakers. The latter prefer to use *somewhat* to modify adjectives, while Chinese English learners prefer to use it to modify verbs for lowering the degree of the action. The cause of this difference lies in the differences of the two languages. According to Lian Shuneng (2010), more verbs are used in Chinese which makes Chinese dynamic, while native English speakers prefer to use more nouns in their utterance, which makes their language static. For the Chinese English learners, they are unconscious of this striking difference between the two languages and transferring the features of their native language to the target language, which leads to the gap in the collocates of *somewhat*.

3. Colligation

In the written language sub-corpus of BNC, there are altogether 4317 concordance lines with *somewhat* as the node. In the present study, 86 samples are further retrieved by random sampling (pick out one in every 50 concordance lines). According to these samples, the frequent colligation patterns of the native speakers is summarized as follows: 1. ADV + ADJ (33 items like *different, harsh, liberal, revolutionary*) 2. ADV + ADJ + N (28 items like *incautious nature, dull performance, mystical feature*) 3. ADV + PREP (10 items like *to my surprise, to Henry's disappointment, after its initial wartime*) 4. ADV + ADV (10 items like *nervously, severely, surprisingly, unceremoniously*) 5. V + ADV (3 items like *recovered, withered*) 6. ADV + V (two items: *manipulate, undermine*)

The frequent colligation of *somewhat* in CLEC are: a. ADV + V (five items: *influence, appears, altered, ruined, lessen*) b. ADV + ADJ (four items: *difficult, unsuitable, hopeful, compulsory*) c. ADV + N (two items: *criminals, a kind of proof*) d. ADV + ADJ + N (one item: *Stygian staircase*) e. ADV + PREP (two items, *like*) f. ADV + SENTENCE (one item: *we are more busy than before.*)

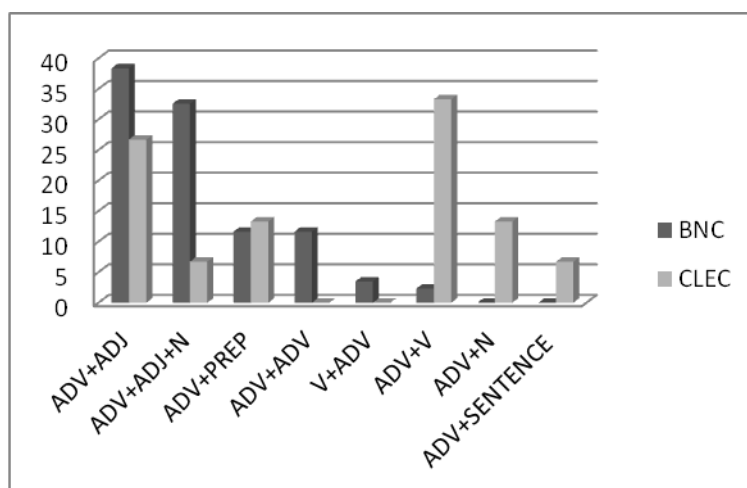


Figure 1: Colligation of *somewhat* in BNC and CLEC

Figure 1 indicates the colligation of *somewhat* in BNC and CLEC. The colligation of the Chinese English learners has the following features compared with the native speakers: a. use less ADV+ADJ and ADV+ADJ+N; b. use more ADV+V; c. misuse ADV+N and ADV+SENTENCE which are never used by the native speakers.

The study shows that the Chinese English learners have mastered the major function of *somewhat*—using it as downtoner to modify adjective and lower the degree and reduce the tone of this word. The colligations used in CLEC are greatly deviant from those in BNC. The Chinese English learners created some colligations that do not conform to the language norms of the native speaker. Most of the Chinese English learners think that learning English means “vocabulary + module test”, while neglecting the acquisition of the colligation of English words. As to the vocabulary, most of them only endeavor to memorize the spelling of each word, while neglecting the collocation and colligation of the word. In using English, they prefer the pattern of “blank-filling”, that is, they will first establish the sentence structure in Chinese and then finish the sentence by filling the blanks by using the English words. In choosing the appropriate word, their first consideration is the Chinese meaning of the word but not native speakers’ habit of using that word. That is why some students even use *somewhat* in some ways that are deviant from those of the native speakers.

4. Semantic Preference

In BNC, the frequent collocations of *somewhat* can be classified into two groups:

a. The words that are used to describe the nature or the relationship of certain things. For example, there are words describing complexity as *complicated/ simplistic/ abstract/ variable/ chequered/ superficial*, and words describing similarity and discrepancy as *different/ similar /differently/ akin*. Biber et al (2000) pointed out that the function of adverbs of degree in academic writing is to express the degree of discrepancy. The present study shows the words expressing similarity and discrepancy are not only the significant collocations of *somewhat* in academic writing but also the collocations in all written language.

b. The words that are used to describe the emotion or attitude. As words to describe surprise like ‘*surprising/ surprised*’; words to describe doubt and uncertainty like ‘*confused/ vague/ ambiguous/ uncertain/ misleading/ confusing/ dubious/ obscure/ blurred/ controversial/ illusory/ shaky/ bewildered/ ambivalent/ bemused/ suspect/ skeptical*’; and words describing unwillingness like ‘*grudging/ reluctant*’.

Downtoners has a lowering effect on the meaning and tone of the language and the use of downtoners can avoid expressions that are too absolute. Both the study of Maocheng Liang (2003) and Xia Zhang indicate that Chinese English learners overuse some intensifiers. This result indicates that Chinese English learners try to avoid plain expression by using some intensifiers, but most of them prefer the use of amplifiers with *very* as an example. The use of

these intensifiers can upgrade the mood and meaning of the language and leave a deep impression on the readers (most of them are the exam markers). The use of downtoners can avoid some expressions that are too exaggerating and absolute, thus making the language mild and negotiable. Chinese English learners do not conform to the language rules of the native speakers in using *somewhat*, which cause the divergence of the semantic preference of this word.

5. Semantic Prosody

The significant collocations of *somewhat* in BNC take on negative meaning like ‘misleading’ and ‘uncertain’, as a result, the semantic prosody of *somewhat* is negative. On the contrary, the semantic prosody of *somewhat* in CLEC is mixed with both collocations expressing positive meaning like ‘hopeful’ and collocations expressing negative meaning like ‘ruin’ and ‘unsuitable’. This discrepancy is largely due to the over-emphasis of the memorizing of the spelling and Chinese meaning of the English words in the English teaching and learning of China. The trend of over-emphasizing the spelling and meaning of the English word lead to the neglect of the collocation, colligation and the semantic prosody and the word. It’s advisable for the English teachers to put the language learning in a broader framework that is the extended meaning unit and put more emphasis on the observation and acquisition of the semantic prosody of each word.

6. Development Pattern

The present study shows that the Chinese English learners acquisition of *somewhat* is gradable and progressive. As to the usage of *somewhat*, the primary English learners cannot use it, the medium English learners use more and the advanced English learners use the most. The result indicates that with the improvement of the learners’ language proficiency, the frequency of the use of *somewhat* develops gradually, which is clearly shown in figure 2. But there is no significant difference in the use of *somewhat* between st4 and st3 in the medium English learners, and also there is no significant difference between st6 and st5 in the advanced English learners. This result shows that during a certain period of the learning process, there will be acquisition plateau. In the acquisition plateau, the learners’ use of specific language feature does not develop with the improvement of their language proficiency, or it is fossilized. Figure 3 shows the gradable feature of the Chinese English learners’ usage of *somewhat*.

According to the features of the development pattern of the Chinese English learners shown in this research, the Chinese English teachers should try to arouse the language learners’ awareness of learning English vocabulary in a broader language context and try to assist them to imitate the native speakers’ way of using the language. In this way, the Chinese English learners can avoid transferring the language features of their first language and truly master the language like a native speaker. On the other hand, the fossilization stage, which is a great hamper in language learning, can be greatly shortened by the constant input of “real English”.

TABLE 4
FREQUENCY OF SOMEWHAT IN EACH SUB-CORPUS

Sub-corpus	Raw frequency	Standard frequency (per million)
St2	0	0
St3	2	9.6113
St4	2	9.3961
St5	5	23.3089
St6	6	26.5362
Sum	15	14.0108

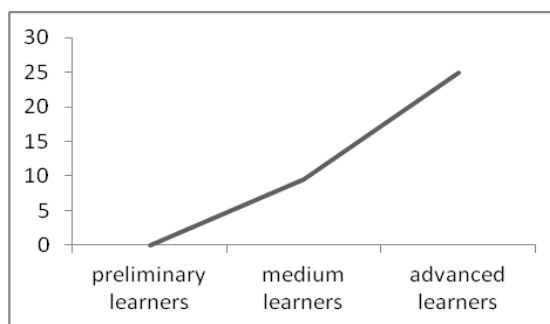


Figure 2: development pattern of Chinese English Learner's usage of somewhat

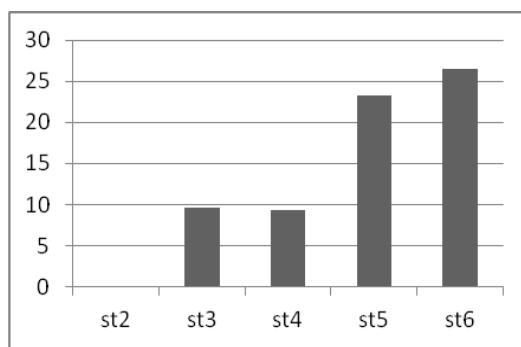


Figure 3: frequency of somewhat in the sub-corpus of CLEC

IV. CONCLUSION

The appropriate use of downtoners, which has a lowering effect in the tone and meaning of the language, can avoid mechanical and inflexible language usage. The Chinese English learners use less *somewhat* compared with the native English learners. The underuse of the downtoners can cause the confusion of the degrees of language expression. Besides the frequency, the usage of Chinese English learners differs greatly in the significant collocations, colligation, semantic preference and semantic prosody. The cause of this discrepancy lies in the inappropriate language teaching and learning habit. English teachers in China prefer to teach and explain the word in isolation, while neglecting the acquisition from in a broader framework—learning the word in the framework of extended meaning unit. The students form a bad view of treating memorizing words as the equivalent of learning English. In this sense, it is advisable to carry out the vocabulary teaching under the framework of extended meaning unit to make the students know not only the surface meaning of the English vocabulary but also the deep information of the vocabulary, including: frequent collocations, the grammatical collocations and its pragmatic function. Moreover, the corpus-driven teaching method can be employed to stimulate the interests of English learners and make them acquire the language in real language data.

REFERENCES

- [1] Biber, D. et al. (2000). *Longman Grammar of Spoken and Written English*. Beijing: Foreign Language Teaching and Research Press.
- [2] Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and Formulae in. A.P. Cowie(Eds) *Phraseology: Theory, Analysis and Applications*. Oxford: Oxford University.145-160.
- [3] Kennedy, G. (2003). Amplifier collocations in the British national corpus: implications for English language teaching. *TESOL Quarterly*, 37: 467–487.
- [4] Lian ShuNeng. (2010). *Contrastive Studies of English and Chinese*. Beijing: Higher Education Press.
- [5] Liang MaoCheng. (2003). A Corpus-based Study of Intensifiers in Chinese EFL Learners' Oral Production. *Asian Journal of English Language Teaching*, 14:105-108.
- [6] Lorenz, G. (1998). Overstatement in advanced learners' writing stylistic aspects of adjective intensification. *Learner English on Computer*. London: Addison Wesley Longman.53-66.
- [7] Quirk, R. (1985). *A Comprehensive Grammar of the English Language*. London and New York: Longman.
- [8] Sinclair, J. (2004). New evidence, new priorities, new attitudes. In Sinclair, J. (Eds) *How to Use Corpora in Language Teaching*. Amsterdam: John Benjamins.
- [9] Sinclair, J. & R. Carter. (2004). *Trust the Text Language, Corpora and Discourse*. London: Routledge, 2004.
- [10] Zhang Xia. (2005). Study on the Use of Intensifiers of Chinese non-English EFL learners—A Research based on COLSEC. In Wei Naixing (Eds) *Corpora in Use*. Shanghai: Shanghai Foreign Language Education Press, 178-194.

Qian Yang was born in Zibo, China in 1981. She received her master degree in Linguistics from Shandong Normal University, China in 2006. She is currently a lecturer in the school of Applied English studies Shandong University of Finance and Economics, Jinan, China. Her research interests include corpus linguistics, translation and cross-cultural studies.