



## A Corpus-based Study of Developmental Stages of Demonstratives in Chinese English Majors' Writing

Jie Sun (Corresponding author)

School of Foreign Languages, Ludong University  
186 Hong Qi Zhong Street, Yantai 264025, Shandong, China  
Tel: 86-535-668-3061 E-mail: sun1978jie@163.com

Xia Liang

No. 1 Middle School  
An Dong Wei Street, Lanshan District, Rizhao 276800, China

*The research is financed by Ludong University. No. W20072601*

### Abstract

On the basis of comparisons between the corpus of English Majors' Composition of Ludong University (EMC corpus) and other corpora as well as among different levels in EMC corpus, this thesis studies the feature of using frequency and stages of development of demonstratives used by Chinese English Majors. The result reveals that English Majors of lower and higher grades tend to overuse and underuse demonstratives respectively; with their better and better command of English, the using frequency of demonstratives tends to be closer to that of native speakers' except for the proportion of the using frequency of singular to plural demonstratives as demonstrative pronouns as well as the using frequency of plural demonstratives.

**Keywords:** Demonstratives, Frequency, Developmental stages, Features

### 1. Introduction

Demonstratives play an important role in English writing. Appropriate use of demonstratives can make the discourse coherent and structure compact (Zhang Yan, 2006).

Demonstratives consist of demonstrative pronouns and demonstrative determiners. A demonstrative pronoun can act as anaphora and cataphora, being used instead of a noun to designate an object which is identifiable from context or usage. A demonstrative determiner is a modifying word which limits the application or reference of the noun modified. Anaphora is the use of an expression which refers to or stands for an earlier word or group of words. Cataphora is the use of an expression which refers to or stands for a later word or group of words (*Oxford English Dictionary*, 2004: 75-2366).

Many scholars have conducted researches about demonstratives. Zhang Ji (2003) studies demonstratives as anaphora and cataphora, yet the study lacks the support of data from corpora. Chen Yi (2002) finds that the frequency of "that" is much higher than "this" with the statistical sampling from the British and American English vocabulary, but the sample is about native speakers which can not reveal the situation of Chinese English learners. Although Zhang Yan (2006) makes a comparison with native speakers on the basis of WECCL corpus and reveals features of demonstratives (this, these, that and those) used by Chinese English learners, her corpus is static and unable to reveal Chinese English learners' developmental stages. So this thesis, on the basis of the corpus of English Majors' Compositions of Ludong University (EMC corpus), carries out a research concerning features of frequency distribution in Chinese English Majors' using demonstratives as well as features at different developmental stages through comparing with other corpora and between different grades in this corpus. The result of this study can give writing instruction some enlightenment.

## 2. Research design

### 2.1 Objectives of the research

- 1) What are the features of the frequency of demonstratives used by Chinese English Majors?
- 2) What are the developmental features of the frequency of demonstratives at different stages?

### 2.2 Corpus

The corpus consists of compositions of English Majors in Ludong University. The compositions originate from the final exam on comprehensive English of Grade 2008, 2007 and 2006 respectively in the first semester of the school year 2008-2009. The token of the corpus is 87,606 words, with 31,348 words in Grade 2008, 30,381 words in Grade 2007, and 25,877 words in Grade 2006.

### 2.3 Method and procedure

In the first place, compositions are input into computers manually and saved as documents (.txt), and after a random sampling check, the accurate rate of the input is 100 percent. Second, with the help of the software "Concord" and "Wordlist" in "Wordsmith", examples containing the four demonstratives (this, these, that and those) are collected, and then examples with "that" serving as the relative pronoun and adverb are deleted manually. Third, all the rest examples are classified in accordance with the demonstrative pronoun, demonstrative determiner, anaphora and cataphora manually. Finally, frequency in the corpus and each grade is calculated and features are revealed.

## 3. Data analysis and discussion of EMC Corpus

### 3.1 Analysis of frequency distribution of demonstrative pronouns and demonstrative determiners

Table 1 is about learners' using frequency of demonstratives in EMC corpus. The figures after each demonstrative are the times that these demonstratives appear per 10,000 words on average in the corpus.

#### Insert Table 1 Here

Table 1 tells that:

First of all, concerning the overall using frequency of four demonstratives, "this" (36.3) is most frequently used; "that" (17.05) and "these" (7.3) are used less; "those" (3.76) is used the least.

Secondly, in terms of the using frequency of demonstrative pronouns and demonstrative determiners, the overall using frequency of demonstrative pronouns is much less than that of demonstrative determiners (22.18<42.23). Specifically speaking, the using frequency of "this", "these" and "those" as demonstrative pronoun is less than that of each corresponding demonstrative determiner (9.02<27.28, 2.28<5.02, 0.68<3.08), while the using frequency of "that" as the demonstrative pronoun is higher than the demonstrative determiner (10.2>6.85). The column of "Pro" shows "that" (10.2) is used most frequently. The next is "this" (9.02), and then "these (2.28)"; "those (0.68)" is used the least. The column of "Det+N" reveals that "this" (27.28) is most frequently used, with "that (6.85)", "these (5.02)" next to it, and "those (3.08)" has the lowest using frequency.

Thirdly, as for the using frequency of singular and plural demonstratives, the overall using frequency of singular demonstratives "this" and "that" is higher than plural ones (53.35>11.06). As demonstrative pronouns and demonstrative determiners, singular ones are both more frequently used than plural ones (19.22>2.96, 34.13>8.1).

The frequency distribution of singular demonstratives is as follows: The overall using frequency of "this" is higher than "that" (36.3>17.05). Being demonstratives pronoun, "this" is less used than "that" (9.02<10.2), but more as demonstrative determiner (27.28>6.85). As for "this" itself, it is more applied as determiners than pronouns (27.28>9.02); the using condition of "that" is opposite (6.85<10.2).

The frequency distribution of plural demonstratives is as follows: the overall using frequency of "these" is higher than "those" (7.3>3.76); "these" is more used than "those" both as pronouns (2.28>0.68) and as determiners (5.02>3.08); as pronouns, the using frequency of "these" is lower than that of determiners (2.28<5.02), and so is "those" (0.68<3.08).

Fourthly, proximal and distal demonstratives are demonstrated in the following part:

The overall using frequency of proximal demonstratives "this" and "these" is more than that of distal demonstratives "that" and "those" (43.6>20.81). As both demonstrative pronouns and demonstrative determiners, the frequency of proximal demonstratives is higher than that of distal demonstratives (11.3>10.88, 32.3>9.93).

As to the two proximal demonstratives, "this" is more preferred than "these" in not only the total using frequency (36.3>7.3), but also the frequency of being pronouns (9.02>2.28) and determiners (27.28>5.02).

As for the two distal demonstratives, "that" is used more often than "those" in overall frequency (17.05>3.76) in being demonstrative pronouns (10.2>0.68) as well as demonstrative determiners (6.85>3.08).

Table 1 reveals that learners in EMC corpus prefer the form “Det+N” to “Pro” (42.23>22.18). What’s more, they would rather use singular demonstratives “this” and “that” than plural “these” and “those” (53.35>11.06). They would rather use proximal demonstratives “this” and “these” than distal “that” and “those” (43.6>20.81). At the same time, learners in EMC corpus possess two more distinctive features when they use the form of “Det+N”: firstly, they more often repeat words that have been mentioned before such as “passage, question, essay, text, people, factors, reasons, case, term, ability, condition, thought, competition, situation, etc”, which do not convey new meanings, and can not be added explanations and complements to. Secondly, they usually use lots of general words without workable implication, like “way, thing, time, moment, aspect, level, concept, year, etc”.

### 3.2 Analysis of frequency distribution of anaphoric and cataphoric demonstratives

Table 2 is learners’ frequency distribution of anaphoric demonstratives and cataphoric demonstratives in EMC corpus. The figures in column “Anaphora” are the percentage resulting from dividing all the times of each demonstrative by the times of each corresponding demonstrative being anaphoric demonstrative. The figures in column “Cataphora” equal 100 percentage minus the percentage in the corresponding column of “Anaphora”.

#### Insert Table 2 Here

According to table 2, in general, most demonstrative pronouns work as anaphora whereas only a minor part acts as cataphora (94.33%>5.67%). The column of “Anaphora” shows the highest using frequency comes from “this” and “these” (100%), and then “that (93.26%)”; “those (16.67%)” is the lowest. The column of “Cataphora”, however, indicates “those (83.33%)” is used the most, with “that (6.74%)” next to it. While no “this” or “these” is used as cataphora.

In the point of the using frequency of singular and plural demonstratives, the following features are revealed:

Between singular “this” and “that”, as anaphora, “this” is used more often than “that” (100%>93.26%). *Word Frequencies in British and American English*, however, argues that the using frequency of “that” is far higher than that of “this” as anaphora. This comparison reveals that there is much difference between learners in EMC corpus and native speakers when “this” and “that” act as anaphora. While working as cataphora, “that” is used more frequently than “this” (6.74%>0). But Chen Ping (1994) proves that “this” is usually chosen to be cataphora, whereas “that” is occasionally used in this way. This comparison tells that learners in EMC corpus need to improve the use of singular demonstratives as cataphora. Results from the former comparison indicate that it is because of misuse and wrong use of singular demonstratives that leads to a much different result with the proper use of these words. Therefore, English teachings in the future should pay more attention to this aspect so as to guide learners and ensure that they can use them correctly.

Between plural “these” and “those”, as anaphora, “these” is used much more frequently than “those” (100%>16.67%), while as cataphora, it’s the opposite (0<83.33%). Yu Hongrong (1998) points out if demonstratives are used as cataphora, people would prefer proximal pronouns. Therefore, learners in EMC corpus still need to better the use of plural demonstratives as cataphora.

In the aspect of the using frequency of proximal and distal demonstratives, the characteristics are:

As for proximal demonstratives’ being anaphora, the using frequency of “this” and “these” gets to 100%, while being cataphora, their occurrence is zero. According to Zhang Ji (2003), “this” and “these” can occasionally be used as cataphora, referring to a whole sentence or sentences of the next part of the text. As a result, it seems that learners in EMC corpus have a good learning of proximal demonstratives as anaphora. However, there are no examples in E corpus that “this” and “these” indicate a whole sentence or sentences, so the case of working as cataphora needs to be proved with a larger corpus.

Concerning distal demonstratives “that” and “those”, the using frequency of “that” is much higher than that of “those” as anaphora (93.26%>16.67%), whereas it is much lower than that of “those” as cataphora (6.74%<83.33%). In other words, “that” is most probably used as anaphora while “those” is very likely to be used as cataphora. Whereas Zhang Ji (2003) points out that “that” and “those” seldom appear when serving as cataphora. If they are used occasionally, they probably convey the sense of irony. Thus learners in EMC corpus need to take further steps to learn the usage of “those” as cataphora.

All in all, learners of EMC corpus use more anaphoric demonstratives than cataphoric ones. From the point of singular and plural demonstratives, there is a wide disparity in using singular demonstratives and plural demonstrative “these” being cataphora needs further improvement. From the point of proximal and distal demonstratives, the acquisition of proximal demonstratives as anaphora is relatively good, but the acquisition of the distal demonstrative “those” needs improving.

#### 4. Comparison between different corpora and between different grades

##### 4.1 Comparison of frequency distribution of demonstratives between different corpora

Table 3 displays the frequency distribution of demonstratives in different corpora. Results from corpora WECCL (W) and LOCNESS (L) (ZhangYan, 2006) are also listed to make an inter-corpus comparison. Results of corpora WECCL (W) and LOCNESS (L) are from ZhangYan (2006). Figures in the table are the times that these demonstratives appear per 10,000 words on average in each corresponding corpus.

##### Insert Table 3 Here

Table 3 reveals the following characteristics:

To begin with, the using frequency of demonstratives in EMC corpus is much higher than that of both natives (L) and W corpus ( $64.37 > 6.89 > 3.97$ ). Because learners in W corpus are Chinese English majors of Grade Three and Four, and learners in L corpus are junior and senior native speakers in American University. So the conclusion might be drawn that: compared with native speakers, Chinese English majors of lower grades have the tendency of overusing demonstratives, whereas those of higher grades have the trend to underusing them.

In the second place, characteristics of the using frequency of demonstrative pronouns and demonstrative determiners are: (1) In the respect of the using frequency of demonstrative pronouns, learners in EMC corpus have a higher frequency than native speakers' and learners in W corpus ( $22.14 > 2.74 > 1.47$ ), and so is the situation in the respect of the using frequency of demonstrative determiners ( $42.23 > 4.15 > 2.5$ ). It seems that compared with native speakers, Chinese English majors of lower grades have the tendency to overusing both demonstrative pronouns and demonstrative determiners whereas those of higher grades have the trend to underusing them. (2) In terms of the using frequency between demonstrative pronouns and demonstrative determiners, the using frequency of demonstrative determiners is higher than that of demonstrative pronouns in each corpus (EMC/W/L) ( $42.23 > 22.14$ ,  $2.5 > 1.47$ ,  $4.15 > 2.74$ ). This seems to lead to a conclusion that even though Chinese English majors (EMC&W) have the tendency to overusing or underusing demonstratives, they obtain similar proportion between the using frequency of demonstrative pronouns and determiners with native speakers, that is, the using frequency of demonstrative pronouns is higher than that of demonstrative determiners.

In the third place, in the aspect of the using frequency of singular and plural demonstratives, characteristics are: (1) The overall frequency of singular demonstratives is higher than that of plural ones in each corpus (EMC/W/L) ( $53.31 > 11.06$ ,  $2.75 > 1.22$ ,  $5.26 > 1.63$ ). (2) As to the using frequency of demonstrative pronouns, the singular form is more often used than the plural form in each corpus (EMC/W/L) ( $19.18 > 2.96$ ,  $1.22 > 0.25$ ,  $2.41 > 0.33$ ). (3) As for the using frequency of demonstrative determiner, the overall using frequency of singular form is higher than plural form in each corpus (EMC/W/L) ( $34.13 > 8.1$ ,  $1.53 > 0.97$ ,  $2.85 > 1.3$ ). Obviously, like native speakers, Chinese English majors (EMC & W) incline to using singular demonstratives more frequently than plural ones.

The situation between singular demonstratives "this" and "that" in each corpus (EMC/W/L) is that: (1) "this" is more often employed than "that" ( $36.3 > 17.01$ ,  $1.89 > 0.86$ ,  $4.82 > 0.44$ ); (2) Chinese English majors (EMC & W) prefer to use "that" rather than "this" as pronouns ( $10.16 > 9.02$ ,  $0.66 > 0.56$ ), whereas the case of native speakers is not so ( $0.22 < 2.19$ ); (3) both Chinese English majors (EMC & W) and natives (L), would rather use "this" than "that" as determiners ( $27.28 > 6.85$ ,  $1.33 > 0.2$ ,  $2.63 > 0.22$ ). All these show that Chinese English majors (EMC & W) have the similar tendency with native speakers in the relative proportion between "this" and "that" in terms of overall frequency as well as the frequency of being determiners, whereas the frequency of being pronouns is opposite.

The situation between plural demonstratives "these" and "those" in each corpus (EMC/W/L) is that: (1) In the respect of overall using frequency, "these" is more frequently used than "those" ( $7.3 > 3.76$ ,  $0.66 > 0.56$ ,  $1.25 > 0.38$ ). (2) Chinese English majors (EMC & W) prefer to use "these" rather than use "those" as pronouns ( $2.28 > 0.68$ ,  $0.15 > 0.1$ ), while the case of native speakers is opposite ( $0.11 < 0.22$ ). (3) When the two demonstratives work as determiners, the use of "these" is more often than "those" in each corpus ( $5.02 > 3.08$ ,  $0.51 > 0.46$ ,  $1.14 > 0.16$ ). From the above, conclusions can be revealed that Chinese English majors (EMC & W) tend to use "these" and "those" similarly with native speakers in the relative proportion between them in terms of overall frequency as well as the frequency of being determiners, whereas the frequency of being pronouns is opposite.

In the fourth place, in the point of the using frequency of proximal and distal demonstratives, characteristics are: (1) People favor to employ more proximal demonstratives than distal ones in each corpus (EMC/W/L) ( $43.6 > 20.77$ ,  $2.55 > 1.42$ ,  $6.07 > 0.82$ ). (2) In terms of demonstrative pronouns, figures in EMC and L corpus show the using frequency of proximal demonstratives is higher than that of distal ones ( $11.3 > 10.84$ ,  $2.3 > 0.44$ ), whereas figures in W corpus indicate the opposite ( $0.71 < 0.76$ ). (3) Proximal demonstratives are more favored than distal ones when they work as determiners. It is striking to see that Chinese English majors (EMC & W) have the same tendency with native speakers in the relative proportion between proximal and distal demonstratives in terms of overall frequency, the frequency of being determiners as well as the frequency of being pronouns in EMC corpus.

The frequency of proximal demonstratives “this” and “these” in each corpus (EMC/W/L) is that: (1) the using frequency of “this” is higher than that of “those” ( $36.3 > 7.3$ ,  $1.89 > 0.66$ ,  $4.82 > 1.25$ ); (2) the same frequency distribution works on their being pronouns and determiners ( $9.02 > 2.28$ ,  $0.56 > 0.15$ ,  $2.19 > 0.11$ ;  $27.28 > 5.02$ ,  $1.33 > 0.51$ ,  $2.63 > 1.14$ ). The comparisons can be summarized as: in the relative proportion of proximal demonstratives “this” and “these” in terms of overall frequency, the frequency of being pronouns and determiners, Chinese English majors (EMC & W) have similar results with native speakers.

The frequency distribution of distal demonstratives “that” and “those” in each corpus (EMC/W/L) is that: (1) the using frequency of “that” is higher than that of “those” ( $17.01 > 3.76$ ,  $0.86 > 0.56$ ,  $0.44 > 0.38$ ); (2) Chinese English majors (EMC & W) incline to using “that” more than “those” as pronouns ( $10.16 > 0.68$ ,  $0.66 > 0.1$ ), but native speakers use the two with the same frequency ( $0.22$ ); (3) “that” is more frequently used than “those” in EMC corpus and L corpus as determiners ( $6.85 > 3.08$ ,  $0.22 > 0.16$ ), while the case is opposite in W corpus ( $0.2 < 0.46$ ). As a result, the conclusion can be drawn that learners in EMC corpus tend to have similarity with native speakers in the relative proportion between “that” and “those” in terms of the overall frequency as well as the frequency of being determiners. To be more concrete, “that” is applied more frequently than “those”, and there is still some room for improving the proportion between “that” and “those” when Chinese English majors (EMC & W) use distal demonstratives as pronouns.

Lastly, with native speakers’ using frequency being the criteria, conclusions can be summarized. Even though learners in EMC corpus incline to overusing demonstratives while learners in W corpus have the opposite tendency, they share similar tendencies with native speakers in the relative proportion between demonstrative pronouns and demonstrative determiners in terms of overall frequency, between singular and plural demonstratives in terms of overall frequency as well as the frequency of being pronouns and determiners, between singular demonstratives “this” and “that”, plural demonstratives “these” and “those” in terms of overall frequency as well as the frequency of being determiners, between proximal and distal demonstratives in terms of overall frequency, the frequency of being determiners as well as the frequency of being pronouns in EMC corpus, between proximal demonstratives “this” and “these” in terms of overall frequency as well as the frequency of being pronouns and determiners, between distal demonstratives “that” and “those” in terms of overall frequency as well as the frequency of being determiners. Whereas they share opposite tendencies with native speakers in the relative proportion between singular demonstratives “this” and “that”, plural demonstratives “these” and “those” as well as distal demonstratives “that” and “those” as pronouns. So the conclusion is that although learners in EMC corpus are inclined to use the form of “Det+N” as native speakers, there are also some differences in specific usage. Most of the nouns used by native speakers express the illocutionary act and psychological process such as “claim, statement, viewpoint, quote, argument etc”. These words’ implication can supply new information, link and develop discourse (Zhang Yan, 2006). While learners in E corpus mostly use general words without much implication, like “way, thing, time, moment, aspect, level, concept and year etc”. On the one hand, this phenomenon shows that the vocabulary of learners in EMC corpus is limited, and only common words are used frequently; on the other hand, it also indicates that learners in EMC corpus only consider grammar and ignore the importance of diction to the coherence of discourse when they write. So vocabulary and instruction of discourse should be paid attention to in the future.

#### 4.2 Comparison of frequency distribution of demonstratives between different grades

##### 4.2.1 Comparison of frequency distribution of demonstrative pronouns and demonstrative determiners

Table 4 is the frequency distribution of demonstratives in EMC corpus, ranging from Grade One to Grade Three, in which “G1” represents Grade One, and “G2”, “G3” are Grade Two and Grade Three respectively. Figures in each column of each grade represent the times that these demonstratives appear per 10,000 words on average in the corpus of each grade. For the sake of convenient comparison the using frequency of native speakers (L) is listed as well, and the figures of L corpus are taken from Table 3.

#### Insert Table 4 Here

What can be found in table 4 is that:

First, the overall frequency of demonstratives from Grade One to Three is far higher than that of native speakers’ ( $52.95 > 6.9$ ,  $77.76 > 6.9$ ,  $52.83 > 6.9$ ). It seems to indicate that compared with native speakers, all students in the three grades, especially Grade Two, tend to overuse demonstratives, and Grade Three inclines to being the closest to the native. It is probable because one semester’s learning of the course of grammar in Grade Two causes students not to have a good command of the newly-learned knowledge. With the broadening amount of new knowledge, they can not absorb the newly-learned information in time, so problems in using demonstratives appear. However, a second semester’s learning of grammar together with one more semester’s consolidation in Grade Three leads learners to progress, so when they are in Grade Three, their using frequency is the closest to natives’.

Second, as for the using frequency of demonstrative pronouns and demonstrative determiners, students from Grade One to Three prefer to use more demonstrative determiners than demonstrative pronouns ( $29.66 > 23.29$ ,  $54.73 > 23.03$ ,

35.27>17.64), and so do native speakers (4.16>2.74). The using frequency of demonstrative pronouns in the three grades is much higher than that of native speakers', however, the higher the grade is, the lower the using frequency appears, and the closer frequency is to the natives'. It proves that the higher the learners' English proficiency is, the closer the frequency is to native speakers'. Demonstrative determiners are used more often by Chinese English majors: Grade Two is the most distinctive (54.73>4.16), with Grade Three the second (35.27>4.16), and Grade One the last (29.66>4.16). It seems to indicate that the learners' language level has no certain connection with the using frequency of demonstrative determiners.

Third, in the aspect of the using frequency of singular and plural demonstratives, characteristics are as follows:

From Grade One to Three the overall using frequency of singular demonstratives "this" and "that" is higher than that of the corresponding plural demonstratives "these" and "those" (46.25>6.7, 65.14>12.62, 37.16>15.75), with the proportion of 6.9, 5.16 and 2.36. In other words, with the improvement of the learners' language level, the proportion decreases successively, which is nearer and nearer to native speakers' 3.21.

In the respect of the using frequency of demonstrative pronouns and demonstrative determiners, characteristics are as follows: (1) from Grade One to Three, the proportion of the overall frequency of singular demonstrative pronouns to the overall frequency of plural demonstrative pronouns is 13.56, 5.66 and 3 respectively, which takes on the decreasing tendency and tends to be lower than natives' 7.3. As singular demonstrative pronouns, the overall using frequency of freshmen is the highest (21.69), with sophomores (19.58) and juniors (13.23) next to it. To be more exact, with the improvement of the learners' language level, singular demonstrative pronouns are used less and less, and thus closer to native speakers' 2.41. However, the using frequency of plural demonstrative pronouns from Grade One to Three is 1.6, 3.46 and 4.41, which becomes higher and higher than natives' 0.33. That is, with the improvement of the learners' language level, the frequency of plural demonstrative pronouns takes a reverse trend. It not only results in the phenomenon that the proportion of the frequency of singular demonstrative pronouns to that of plural demonstrative pronouns becomes lower and lower than native speakers' but also reveals that learners in EMC corpus do a good job in singular demonstrative pronouns, but they need to improve the learning of plural demonstrative pronouns; (2) the proportion of the frequency of singular demonstrative determiners to that of plural demonstrative determiners from Grade One to Three is 4.82, 4.97 and 2.11 respectively, which is closer and closer to that of native speakers' is 2.18.

Fourth, in terms of the using frequency of proximal and distal demonstratives, the following features are discovered: (1) Proximal demonstratives are more frequently used than distal ones from Grade One to Three (31.9>21.05, 52>25.76, 45.35>7.56), which shares the same situation with native speakers (6.08>0.82). The proportion of the overall using frequency of proximal demonstratives to distal ones from Grade One to Three is 1.52, 2.02 and 6 respectively, which shows with the improvement of the learners' language level, the proportion of proximal demonstratives to distal ones becomes increasingly higher and approaches more and more closely to natives' 7.41; (2) as demonstrative pronouns, proximal demonstratives are used less often than distal ones in Grade One (7.66<15.63), whereas it is the opposite in Grade Two and Three (13.37>9.66, 13.23>4.41). What's more, the proportion of the using frequency of proximal demonstrative pronouns to distal ones from Grade One to Three is 0.49, 1.38 and 3, which takes on a rising tendency and is closer and closer to native speakers' 5.23. The using frequency of distal demonstratives of Grade One to Three is 15.63, 9.66 and 4.41, which also moves towards natives' 0.44. In a word, with the improvement of the language level, learners have been making progress in using distal demonstratives; (3) as demonstrative determiners, Chinese English majors of Grade One to Three use proximal demonstrative determiners more often than distal ones (24.24>5.42, 38.63>16.1, 31.49>3.15), which is similar to natives (3.78>0.38). The proportion of the frequency of proximal demonstrative determiners to that of distal demonstrative ones from Grade One to Three is 4.47, 2.4 and 10 respectively. Although the proportion in Grade Two decreases a little, it rises again in Grade Three and becomes the closest to that of native speakers' 9.95.

Fifth, compared with native speakers, the three grades have the tendency of overusing demonstratives, but the overall frequency of Grade Three is approaching the most that of native speakers'. With learners' improvement of the language level, the frequency of demonstrative pronouns, the proportion between singular and plural demonstratives in terms of overall frequency, the frequency of singular demonstratives as demonstrative pronouns, the proportion between proximal and distal demonstratives in terms of overall frequency as well as the frequency of being demonstrative pronouns, the frequency of distal demonstratives as demonstrative pronouns of Grade One to Three are more and more approaching native speakers'. While the proportion between singular and plural demonstratives in terms of frequency as demonstrative pronouns, the frequency of plural demonstratives as demonstrative pronouns and demonstrative determiners are opposite to that of native speakers'.

#### 4.2.2 Comparison of frequency distribution of anaphoric and cataphoric demonstratives

Table 5 is about the frequency distribution of anaphora and cataphora from Grade One to Three. "G1" stands for "Grade One", "G2" "Grade Two" and "G3" "Grade Three". The figures in column "Anaphora" of each grade are the percentage resulting from dividing all the times of each demonstrative by the times of each corresponding demonstrative being

anaphoric demonstrative. The figures in column “Cataphora” of each grade equal 100 percentage minus the percentage in corresponding column of “Anaphora”.

### Insert Table 5 Here

Table 5 reveals that:

Firstly, from Grade One to Three the using frequency of anaphora is much higher than that of the corresponding cataphora (90.41%>9.59%, 96.77%>3.22%, 96.43%>3.57%).

Secondly, concerning the using frequency of singular and plural demonstratives, characteristics are as follows: (1) The singular “this” and “that” are more often used as anaphora than as cataphora in each grade. The frequency of three grades using “this” as anaphora is 100 and 0 as cataphora. As to “that”, when it acts as anaphora, its increasingly rising using frequency of Grade One to Three is 89.36%, 97.22% and 100%; when it acts as cataphora, its decreasing using frequency of Grade One to Three is 10.64%, 2.77% and 0. However, Chen Ping (1994) points out that both “this” and “that” can be used as anaphora, though “that” is more often used. But if they work as cataphora, “this” is used much more often than “that”. So the conclusion can be drawn that students of Grade One to Three in EMC corpus do a gradually better job in the usage of “that” as anaphora and cataphora, but they make no progress in the usage of “this” as cataphora. (2) As for the plural “these” and “those”, “these” is more often used as anaphora, and “those” as cataphora in three grades. When “these” works as anaphora, the frequency of three grades is equal (100%), and as cataphora, its frequency is zero. When “those” is used as anaphora, its using frequency of Grade Two is the highest (33.33%), whereas that of Grade One and Three is 0. As cataphora, the using frequency of Grade One and Three is the highest (100%), the frequency of Grade Two is next to it (66.67%). Yu Hongrong (1998) makes a survey indicating that people incline to using proximal demonstratives as cataphora. There comes out the conclusion: learners in EMC corpus still need to improve the usage of “these” as cataphora.

Thirdly, in terms of the using frequency of proximal and distal demonstratives, characteristics are as follows:

In the respect of proximal demonstratives “this” and “these”, they are used as anaphora in all grades. Zhang Ji (2003) points out that “this” and “these” can also be used as cataphora when they, in the most part, refer to a sentence or some sentences in the following context. In this sense, learners in EMC corpus use “this” and “these” correctly as anaphora. Since the case that “this” and “these” are used as cataphora does not exist, there is no way to expound it.

As for distal demonstratives “that” and “those”, “that” is more often used as anaphora, while “those” as cataphora. When being anaphora, the using frequency of “that” from Grade One to Three is 89.36%, 97.22% and 100%, with an increasingly rising tendency; “those” is favored the most by students in Grade Two (33.33%), while freshmen and juniors never use them. When being cataphora, the using frequency of “that” from Grade One to Three is 10.64%, 2.77 % and 0 with a decreasing tendency; “those” is most frequently used in Grade One and Three (100%), with Grade Two next to it (66.67%). Zhang Ji (2003) indicates that “that” and “those” are seldom used as cataphora. So learners in EMC corpus acquire the usage of “that” as anaphora and cataphora quite well, but not “those”.

Fourthly, all these analyses above lead to the conclusion: With the improvement of the language level, learners make progress in the acquisition of distal and singular demonstrative “that” as anaphora and cataphora. There is no change in the using frequency of proximal demonstratives “this” and “these” as anaphora, which suits for the usage of them. But the learning of singular demonstrative “this” and plural demonstrative “these” as cataphora as well as distal demonstrative “those” as anaphora and cataphora needs improving.

### 5. Conclusion

This research finds that: compared with native speakers, English majors of lower grades in EMC corpus have the tendency of overusing demonstratives, whereas those of higher grades in W corpus have the trend to underusing them. Even though Chinese English majors (EMC&W) have the tendency of overusing and underusing demonstratives, they share similar tendencies with native speakers in the relative proportion between demonstrative pronouns and demonstrative determiners in terms of overall frequency, between singular and plural demonstratives in terms of overall frequency as well as the frequency of being pronouns and determiners, between proximal and distal demonstratives in terms of overall frequency, the frequency of being determiners as well as the frequency of being pronouns in EMC corpus. Whereas they share opposite tendencies with native speakers in the relative proportion between singular demonstratives “this” and “that”, plural demonstratives “these” and “those”, as well as distal demonstratives “that” and “those” in terms of the frequency of being pronouns. At the same time, learners of EMC corpus use more anaphoric demonstratives than cataphoric ones. From the point of singular and plural demonstratives, there is a wide disparity with native speakers in using singular demonstratives, and plural demonstrative “these” as cataphora needs further improvement. From the point of proximal and distal demonstratives, the acquisition of proximal demonstratives as anaphora is relatively good, and the acquisition of the distal demonstrative “those” needs improving.

From the comparison of three grades of EMC corpus, the conclusion can be drawn that compared with native speakers,

three grades have the tendency of overusing demonstratives, but the overall frequency of Grade Three is approaching the most that of native speakers'. With learners' improvement of the language level, the frequency of demonstrative pronouns, the proportion between singular and plural demonstratives in terms of overall frequency, the frequency of singular demonstratives as demonstrative pronouns, the proportion between proximal and distal demonstratives in terms of overall frequency as well as the frequency of being demonstrative pronouns, the frequency of distal demonstratives as demonstrative pronouns from Grade One to Three are more and more approaching native speakers'. While the proportion between singular and plural demonstratives in terms of frequency as demonstrative pronouns, the frequency of plural demonstratives as demonstrative pronouns and demonstrative determiners are opposite to that of native speakers'. At the same time, with the improvement of the language level, learners make progress in the acquisition of distal and singular demonstrative "that" as anaphora and cataphora. There is no change in the using frequency of proximal demonstratives "this" and "these" as anaphora, which follows the usage of them. But the learning of singular demonstrative "this" and plural demonstrative "these" as cataphora as well as distal demonstrative "those" as anaphora and cataphora need to be improved.

The enlightenment of this research to the teaching instruction of English writing concerning demonstratives is that both teachers and learners should attach proper attention to the importance of demonstratives because they have the irreplaceable function in discourse and communication. So in the process of teaching, teachers should pay attention to the explanation of the function and usage of demonstratives, make sure that students can use them properly, and thus promote the development of using demonstratives into a higher level.

This thesis also has some deficiencies. First, this research only involves four demonstratives "this", "these", "that" and "those" and the scale is not comprehensive, so in the future, the number of demonstratives should be enlarged to make a more comprehensive research of them. Second, the data sample is not large enough, a larger sample will be of more help in proving the result of this research.

#### References

- Ariel, M. (1988). Referring and Accessibility. *Journal of Linguistics*, 24, 65 - 87.
- Biber, D. (2000). *Longman Grammar of Spoken and Written English*. Beijing: Foreign Language Teaching and Research Press.
- Chen, Ping. (1994). The anaphora and cataphora of English demonstratives. *Foreign Language Research*, 1, 40-45.
- Chen, Yi. (2002). An analysis of the difference between demonstrative reference in English and Chinese. *Journal of Xiamen Educational Institute*, 2, 5-7.
- Liu, Xiancai. (2002). A multi-perspective analysis of English demonstrative "this" and "that" in contextual reference. *Journal of Shandong Educational Institute*, 6, 83-85.
- McCarthy, M. (1994). It, This and That. In Coulthard, M. (ed.). *Advances in Written Text Analysis*. London: Routledge.
- Oxford English Dictionary*. (2004). Beijing: Foreign Language Teaching and Research Press.
- Petch-Tyson, S. (2002). Demonstratives Expressions in Argumentative Discourse-A Computer Corpus-based Comparison of Non-native and Native English. In Botley, S. & Mcenery, A. M. (eds.). *Corpus-based and Computational Approaches to Discourse Anaphora*. Amsterdam: John Benjamin Publishing Company.
- Wang, Lifei. (2005). *Construction and study of Chinese learners' oral and wirriten English corpus*. Nanjing: Nanjing University.
- Yu, Hongrong. (1998). A comparison of reference function of demonstratives in English and Chinese context and their application in translation. *Foreign Languages Research*, 58, 37-41.
- Zhang, Ji. (2003). A comparison of contextual reference of English and Chinese demonstratives. *Journal of Nanchang College*, 4, 71-73.
- Zhang, Yan. (2006). A corpus-based study of Chinese undergraduates's using demonstratives in argumentative essays. *Foreign Languages and Their Teaching*, 6, 27-29.

Table 1. Frequency distribution of demonstratives in EMC corpus

	Pro	Det+N	Total
This	9.02	27.28	36.3
These	2.28	5.02	7.3
That	10.2	6.85	17.05
Those	0.68	3.08	3.76
Total	22.18	42.23	64.41

Table 2. Frequency distribution of anaphora and cataphora from Grade One to Grade Three in EMC corpus

	Anaphora	Cataphora
This	100	0
These	100	0
That	93.26	6.74
Those	16.67	83.33
Total	94.33	5.67

Table 3. Frequency distribution of demonstratives in different corpora

	Pro	Det+N	Total
	EMC/W/L	EMC /W/L	EMC /W/L
This	9.02/0.56/2.19	27.28 /1.33/2.63	36.3/1.89/4.82
These	2.28 /0.15 /0.11	5.02/0.51/1.14	7.3/0.66 /1.25
That	10.16 /0.66 /0.22	6.85/0.2/0.22	17.01/0.86 /0.44
Those	0.68 /0.1/0.22	3.08 /0.46 /0.16	3.76 /0.56/0.38
Total	22.14/1.47 /2.74	42.23/2.5/4.15	64.37 /3.97/6.89

Table 4. Frequency distribution of demonstratives of different grades in EMC corpus

	Pro	Det+N	Total
	G1/G2/G3/L	G1/G2/G3/ L	G1/G2/G3/ L
This	6.7/10.65/9.45/2.19	22.01/32.69 /23.93/2.63	28.71 /43.34/33.38/4.82
These	0.96/2.72/3.78/0.11	2.23/5.94/8.19 /1.15	3.19/8.66/11.97 /1.26
That	14.99/8.92/3.78/0.22	2.55 /12.88 /0/0.22	17.54 /21.8/3.78 /0.44
Those	0.64/0.74/0.63/0.22	2.87 /3.22/3.15/0.16	3.51 /3.96/3.78 /0.38
Total	23.29/23.03/17.64/2.74	29.66/54.73/35.27 /4.16	52.95/77.76/52.83 /6.9

Table 5. Frequency distribution of anaphora and cataphora from Grade One to Grade Three in EMC corpus

	Anaphora	Cataphora
	G1/ G2/ G3	G1/ G2/ G3
This	100/100/100	0/0/0
These	100/100/100	0/0/0
That	89.36 /97.22/100	10.64/2.77 /0
Those	0/33.33/0	100/66.67/100
Total	90.41/96.77 /96.43	9.59 /3.22/3.57