A Corpus-based Study of Chinese EFL Learners' Employment of *although*

Jingwen Chen[1]

[1] School of International Studies, Sun Yat-sen University, Zhuhai, China

Correspondence: Jingwen Chen, 4-202, St. 9, Tiyuhuayuan Residsence, East Yuncheng Rd., Baiyun District, Guangzhou, Guangdong, China, 510405. Tel: 86-136-3136-9004. E-mail: evechansis@126.com

Abstract

Although is a frequently used subordinating conjunction in English. However, non-nativeness is often observed in Chinese EFL learners' *although* output during pedagogical practice. This paper aims at exploring the characteristics of Chinese EFL learners' *although* employment in Chinese EFL learners' writing. The study is a corpus-based analysis launched under the analytical framework of contrastive interlanguage analysis. The interlanguage hypothesis lays the theoretical foundation of the present study. Texts from two corpora—the Chinese learner English corpus [CLEC] and the "arts and humanities" disciplinary group of the British academic written English corpus [sub-BAWEC]—are analyzed both quantitatively and qualitatively with the help of concordance software Antconc 3.2.1 and statistics program PASW Statistics 18. Based on the findings, conclusions are drawn as follows: 1) Chinese EFL learners tend to underuse *although* and produce mono-structural *although* clauses in their writing. Nevertheless, they share similar preference on deciding *although* placement in clauses with native English speakers; and 2) Factors such as interlingual difference between English and Mandarin Chinese, pedagogical neglect in English classrooms and different cognitive styles influence Chinese EFL learners' *although* employment.

Keywords: *although* employment, corpus-based study, contrastive interlanguage analysis, Chinese EFL learners

1. Introduction

Although is a frequently used English conjunction. Under formal educational system, most Chinese children start learning basic English at the fourth grade of primary school (with the exception that some parents may send their children to English training at an earlier age) and will not start systematic learning of grammatical clauses until middle school. However, as a subordinating conjunction that introduces concessive clauses, *although* is usually taught at a relatively early stage of learning. In several commonly used Chinese English textbook series, *although* appears in the eighth or ninth grade vocabulary lists (see Net. 1 & Net. 2) while its informal counterpart *though* has its earliest appearance in a fifth-grade textbook (see Net. 3). Assumed by many English teachers, it is not difficult for Chinese English learners to master *although*. However, it is very common that non-nativeness of *although* employment appears in the written output of Chinese EFL learners.

As corpus linguistics has been developing rapidly since the 1960s, an increasing number of corpus-based studies have been conducted to examine and explain language phenomena. With the establishment of Chinese EFL learner corpora such as the Chinese learner English corpus established in 2003 and the spoken and written English corpus of Chinese learners established in 2005, complete and detailed studies on Chinese EFL learners' employment of certain lexical items or grammatical constructions become possible. Amid the corpus-based linguistic research trend and motivated by the need to understand the general status of Chinese EFL learners' employment of *although*, the present author intends to conduct a systematic corpus-based study in the hope that the study can summarize the main characteristics of *although* in Chinese EFL learners' output and discuss the possible factors that affect Chinese EFL learners' *although* employment. Considering the space and time limit, only written English output of the Chinese EFL learners is analyzed in this paper.

Based on the aims of the study, the research questions are proposed as follows:

1) What are the characteristics of Chinese EFL learners' employment of *although* in written English compared with native English speakers?

2) What are the possible factors that affect Chinese EFL learners' written output of *although*?

2. An Overview and Existing Studies of *although*

2.1 Descriptions of "although"

As stated by Biber, Johansson, Leech, Conrad and Finegan (1999/2000: 85), *although* belongs to the category of subordinators or subordinating conjunctions in structural English grammar. "Subordinators are words which introduce (mainly finite) dependent clauses. Grammatically, subordinators have a purely syntactic role, and this distinguishes them from other clause initiators" (ibid.). Specifically, *although* is one of the subordinators which introduce adverbial clauses of concession, also known as concessive clauses. As *Collins COBUILD English Grammar* explains, people will utter a concessive clause when they "want to make two statements, one of which contrasts with the other or makes it seem surprising" (1990:264). Despite Biber et al.'s argument of subordinators introducing mainly finite dependent clauses, *although* can be used in non-finite concessive clauses. For example, people can say *although liking coffee, I never drank too much* as well as *although I like coffee, I never drank too much* (c.f. *Collins COBUILD English Grammar*, 1990: 265). Moreover, *although* can be used prior to noun groups, adjective groups and adjuncts. All the following sentences are correct:

Although fond of Gregory, she did not love him.

It was an unequal marriage, although a stable and long-lasting one.

They agreed to his proposal, although with many reservations (ibid.).

Part of the present study explores the range of *although* clause constructions in the written output of both Chinese EFL learners and native English speakers, attempting to investigate whether Chinese EFL learners share similar flexibility in generating different types of *although* clauses with native English speakers.

Other than the concessiveness of *although*, Iten (2000) indicates that *although* has an "adversative" meaning. For example, in *He has long legs although he is a bit short of breath*, the dependent *although* clause possesses a negation of the implication of the main clause. In addition, Quirk, Greenbaum, Leech, and Svartvik (1985) point out that *although* and *though* can connect two clauses with similar situations such as in *Although Sam had told the children a bedtime story, June told them one too (anyway)* (1985: 1099).

Carbonell-Olivares (2009) applies the Theme theory of systemic functional grammar to clarify the discourse function of *although*. When *although* is at the initial placement of a concessive clause, it introduces known/given information in the thematic position, and when *although* appears at the medial placement, it introduces new information in the rhematic position.

2.2 Existing Studies of "although"

Few studies have been conducted concerning the issue of *although*. Most existing studies focus on the distinction between *although* and its informal synonym *though*. Most significantly, Biber et al. (1999) carry out a corpus-based study observing the most common subordinators across registers and dialects. Even though to investigate the status of *although* is not the aim of the study, it reveals that users choose *though* and *although* alternatively to express concessive relationships. As far as the four registers investigated are concerned, findings are that "conversation and fiction show a slightly greater use of *though*, news shows no particular preference", while "in academic prose, *although* is about three times as frequent as *though*" (Biber et al., 1999/2000: 845).

Carbonell-Olivares (2009) conducts a highly comprehensive corpus-based analysis to prove the semantic and pragmatic theoretical description of *although*. The analysis is based on the observation and calculation of data from a specifically compiled corpus of 22 linguistic articles with 100 instances of *although*. Carbonell-Olivares concludes that *although* is most prominently used to express denial of expectation and is "linked intrinsically to the subjectivity of the author" (2009). The finding also indicates that *although* contributes to identifying contrast relation and facilitating its cognitive processing. In addition, Carbonell-Olivares (2009) argues that the employment of *although* in the medial position of a clause can be a rhetorical strategy of the writers to highlight new information that contains the writers' opinion or position.

Studies of *although* in the field of second language acquisition are rare. Spycher (2007) launches a study which discusses how to instruct English learners in American secondary schools to progress in academic writing. Spycher records the progressing process of a student named Ernesto in English academic writing by analyzing his use of *although* to illustrate the effect of appropriate pedagogical instruction.

3. Theoretical Foundation and Analytical Framework

3.1 Interlanguage [IL]

The term "interlanguage" was first adopted by Selinker (1972) to refer to an L2 learner's separate language system which "has a structurally intermediate status between the native and target languages [TLs]" (Brown, 1994/2001:203). Other terms are used by different researchers to indicate the same system. Corder (1971) uses "idiosyncratic dialects" which later he refers to as "language-learner languages" (Corder, 1978) while Nemser (1971) uses "approximative systems". The assumptions under the IL hypothesis are that 1) learner's language is different from both L1 and L2 and possesses its own internal structure; 2) learner's language is within a developing continuum evolving towards the TL; 3) learners at the same level of proficiency will share coincident learner language system in a given contact situation (c.f. Nemser, 1971). Moreover, as Tarone (2001) points out, fossilization, which refers to L2 learners' failure to fully acquire the TL, is a distinct characteristic of IL which accounts for the difference and uniqueness of IL compared to TL. One of the significant objectives of IL research, therefore, is to explain the difference between IL and TL and describe the features and development of IL.

The introduction of the IL hypothesis symbolized the establishment of second language acquisition as an independent academic field. Early researchers such as Selinker, Nemser and Corder laid the foundation of the IL hypothesis while the IL hypothesis has been continuously discussed and revised since its establishment. Tarone (2001) summarizes the major modifications and expansions of this hypothesis. These modifications and expansions mainly facilitate discussions and debates from five perspectives, namely the existence of IL in Children language acquisition, the impact of universal grammar on IL development, the influence of social context on IL development, the possible non-existence of fossilization, and the sociolinguistic communicative competence of IL. In addition, an increasing amount of studies have focused on the cognitive perspective of L2 acquisition and learning which has led to new trends of IL studies since the 1990s. Also for the present study, included in the discussions are the similarity and difference in cognitive process between English and Mandarin Chinese and their influence on the characteristics of Chinese EFL learners' *although* employment.

3.2 Contrastive Interlanguage Analysis [CIA]

The CIA framework provides the analytical tool for the present study. As the traditional contrastive analysis [CA] received increasing doubts and criticism after the 1960s and the studies of IL surged, an attempt at contrastively analyzing native language [NL] and IL emerged at the late 1980s. Learner corpora such as the international corpus of learner English (Granger, 1993; 2003) were under construction to fulfill the need of collecting empirical IL data for contrastive studies. Pery-Woodley (1990) affirms that "the comparison of native and non-native texts is a neglected but promising field". The embryonic statement of CIA first appeared in a contrastive study of connector usage in essay writing of native English speakers and French EFL learners conducted by Granger and Petch-Tyson (1996). Granger and Petch-Tyson explain, "we have called the area of research we are involved in 'contrastive interlanguage analysis' (or CIA) which involves comparing and contrasting what non-native and native speakers of a language do in a comparable situation" (1996).

Granger (1996) later provides a more elaborate definition and explanation of the CIA framework. She emphasizes that CIA is different from the traditional CA because "CIA does not establish comparison between two different languages but between native and learner varieties of one and the same language" (Granger, 1996). CIA largely relies on the utilization of computer learner corpora and consists of two types of comparison: comparison between NL and IL, and comparison of different IL varieties of the same language. A comparable control corpus of NL data is required in the first type apart from the learner corpus. As concluded by Paquot (2010:70), the NL/IL comparison can spot the non-native features in leaner languages while the IL/IL comparison is able to decide whether certain features are peculiar to one IL group or common among different IL groups. Furthermore, during the process of CIA, it is considered significant to investigate contrastive frequencies. Granger (1994, 1996) believes that contrastive frequencies investigations are profitable quantitative comparisons in CIA. The overuse or underuse tendency of the target items can usually be detected by frequency calculations. In addition, concordance software and other computer programs are highly involved in CIA.

Compared to the traditional CA, CIA is superior in the following aspects: 1) CIA can better reveal the nature of the IL because a large quantity of learner data can be collected and analyzed in electronic format with the aid of computers; 2) It is possible to conduct IL investigations on a large scale, for the data processing procedure is automatic or semi-automatic; and 3) CIA can not only analyze learner errors in ILs but also expose non-nativeness of the ILs such as the overuse and underuse situations by analyzing the frequency of specific language items.

CIA is often integrated with the CA pattern in practice. According to Granger (1996), the CA data assist to build

predictions of the IL while the IL data serve as diagnosis against the CA-based assumptions. The practical connections among CIA, CA, corpora and computer technologies are presented in Figure 1.

The present study applies the CIA framework to compare native English speakers and Chinese EFL learners' written output of *although* mainly from three perspectives of the *although* employment. During the process, ILs of different proficiency levels are compared if necessary. The Chinese learner English corpus [CLEC] is the IL database while the "arts and humanities" disciplinary group of the British academic written English corpus [sub-BAWEC] is the control corpus of NL for the present study. Practical support is provided by the AntConc concordance program and the PASW statistics program.
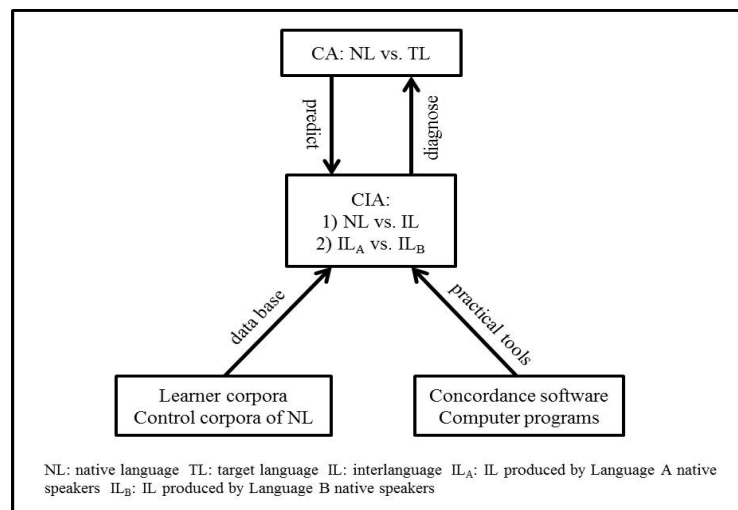

Figure 1. CA, CIA, Corpora and Computer Technologies

4. Data Sources and Methodology

4.1 Data Sources

The present study adopts a corpus-based approach. The data employed in the study are taken from two corpora: CLEC and BAWEC.

CLEC is a learner corpus containing English writing of Chinese EFL learners constructed by Gui and Yang (2003) with 1 207 879 word tokens. Texts in CLEC are collected from daily writing assignments and writing examinations completed by Chinese senior high school students and university students who are considered advanced learners of English. Texts data are further divided into five sub-corpora—ST2, ST3, ST4, ST5 and ST6 in accordance with five English proficiency levels. Each sub-corpus consists of about 20 million word tokens. ST2 contains writing of senior high school students; ST3 includes writing of first-and-second-year non-English majors; ST4 embodies writing of third-and-fourth-year non-English majors; writing of first-and-second-year English majors is included in ST5 while ST6 is made up of writing of third-and-fourth-year English majors. CLEC can represent Chinese EFL advanced learners' IL standard with its text volume and proficiency scale.

BAWEC is a written corpus with in total 6 506 995 tokens. It is the research achievement of a project entitled *An Investigation of Genres of Assessed Writing in British Higher Education* conducted by the universities of Warwick, Reading and Oxford Brookes. It contains 2897 student assignments collected from students who are English native speakers of the above universities in each of three undergraduate years and those in master programs. The assignments were collected from 35 disciplines and in 4 broad disciplinary groupings including art and humanities, life sciences, physical sciences, and social sciences. Only texts from the group of "art and humanities" with 1 450 308 word tokens are selected for the present study as the comparable control corpus considering the proximity of token numbers to CLEC. Moreover, assignments in this group were completed by students majoring in English, history, linguistics and philosophy which makes the student backgrounds and assignment topics similar to those from CLEC. The texts selected for the study are set up as the sub-BAWEC corpus.

4.2 Methodology and Research Procedures

Texts containing *although* are selected and analyzed under the CIA framework. Target texts from both CLEC and sub-BAWEC are compared mainly in terms of frequency, P-value, and chi-square test results. The condition whether *although* is underused or overused is examined with the assistance of the PASW statistic program. By comparing texts from the two corpora, it is expected that the characteristics of Chinese EFL learners' employment of *although* in written English can be summarized and described.

In accordance with one of the basic characteristics of corpus-based studies, the present study conducts both quantitative and qualitative analyses. The research procedures of the study are stated below.

1) Establish the sub-BAWEC corpus.

2) Compare frequency of *although* respectively in CLEC and sub-BAWEC. During this procedure and those follow, chi-square tests are conducted to ensure the results are not chance occurrence.

3) Explore the reasons of the *although* frequency in Chinese EFL learners' written output.

4) Compare respective frequencies in CLEC and sub-BAWEC of the initial and medial placement of *although* in *although* clause constructions.

5) Discuss the *although* placement preference in Chinese EFL learners' *although* clauses output.

6) Calculate respective proportion of different *although* clause construction types in CLEC and sub-BAWEC. (The *although* clause types investigated here are: *although* + finite clause, *although* + non-finite clause, *although* + noun groups, *although* + adjective groups, and *although* + adjuncts.)

7) Compare the variety ranges of *although* clause constructions in CLEC and sub-BAWEC and discuss Chinese EFL learners' *although* clause construction feature.

8) Summarize the findings and describe the characteristics of Chinese EFL learner's *although* employment.

5. Analysis and Discussion

5.1 General Frequencies of "although"

A comparison of frequencies of *although* in the two corpora can reflect the general status of Chinese EFL learners' *although* application in writing. Texts containing *although* are extracted from CLEC and sub-BAWEC by concordance process with AntConc program. Respective raw frequencies of *although* can be obtained from the Wordlist section of the program. Manual check is processed to include cases where *although* is misspelled hence excluded by the program. Standard frequencies with a basis per million word tokens are also normalized from raw frequencies and displayed if necessary. PASW program is utilized to process chi-square tests and results are presented as follows.

5.1.1 *although* Frequency in CLEC and Sub-BAWEC

In mainland China, the conjunction *although* is taught in middle school (junior high school) in most cities. Among the most widely adopted middle school English textbooks, *although* appears generally in the eighth or ninth grade volumes (c.f. Net. 1. & Net. 2.) of the textbook series. Therefore, it is assumed that advanced English learners can generally master the usage of *although* and apply it in writing appropriately. Hence, it is reasonable to predict that there exists no significant difference in general frequencies of *although* in native speakers' and Chinese EFL learners' writing. However, this prediction is denied by the result of quantitative analysis of *although* frequencies in the two corpora.

The general frequencies of *although* in CLEC and Sub-BAWEC are presented in Table 1. As displayed in Table 1, Chinese EFL learners use about 278 *although* per million texts while native English speakers apply about 933 *although* per million texts. The chi-square test result shows a very highly significant difference of general frequencies of *although* (P-value = 0.000 < 0.001) between native English speakers and Chinese EFL learners (Note 1). It is obvious that Chinese EFL learners have a highly significant underuse tendency of *although* in their writing. To further discuss the issue, it is necessary to compare frequencies of *although* in the five sub-CLEC corpora in the attempt to explain why Chinese EFL learners tend to underuse this word in their writing.

Table 1. General frequencies of *although* in CLEC and Sub-BAWEC

Corpus	Total Word Tokens	*Although* Raw Freq.	*Although* Stndd.Freq.(/mil)	X^2	P-value
CLEC	1 207 879	336	278.173 56	444.895 (df=1)	0.000
Sub-BAWEC	1 450 308	1353	932.905 29		

Note. Raw Freq.=Raw Frequency; Stndd.Freq.(/mil) =Standard Frequency per million texts.

5.1.2 *although* Frequencies in Five Sub-CLEC Corpora

As shown in Table 1, 336 cases of *although* employment exist in CLEC with 1 207 879 word tokens. Table 2 presents the frequencies of the 336 *although* hits within the five sub-CLEC corpora.

Table 2. *although* frequencies in five Sub-CLEC Corpora

Corpus	ST2	ST3	ST4	ST5	ST6	CLEC
Frequency	29	31	38	94	144	336
Percentage	8.63%	9.23%	11.31%	27.98%	42.86%	100%

Figures in Table 2 indicate that the employment of *although* follows an ascending trend from ST2 to ST6. Implied by the statistics, senior high school students and university non-English majors (ST2, ST3 and ST4) might be hesitant when employing *although* clauses while university English majors (ST5 and ST6) show more confidence in producing *although* clauses in writing. They produce two and a half times more *although* clauses than non-English majors and senior high school students combined. 42.86% of all *although* employment cases, specifically, are from the ST6 corpus which contains writing of third-and-fourth grade English majors. Generally, Chinese EFL learners seem more confident of applying *although* in their writing as they are reaching higher levels of proficiency.

5.1.3 Chinese EFL Learners' Underuse Tendency of *although*

As observed from the comparative result of *although* employment frequencies between writing texts from CLEC and sub-BAWEC, Chinese EFL learners highly significantly underuse *although* clauses in their written output. And the further comparison among five sub-CLEC corpora indicates an ascending trend of *although* employment from ST2 to ST6. Considering *although* as a formal conjunction with a more informal counterpart *though*, one possible explanation of the underuse tendency of *although* in Chinese EFL learners' writing may lie in Chinese EFL learners' incapability of selecting proper wording for different genres. In order to verify this argument, frequencies of *though* in CLEC and sub-BAWEC are examined and compared, and frequencies of *though* in five sub-CLEC corpora are also calculated to compare with the results in Table 2. The cases where *though* is used as an adverb are excluded manually.

Table 3. General frequencies of *though* in CLEC and sub-BAWEC

Corpus	Total Word Tokens	*Though* Raw Freq.	*Though* Stndd.Freq.(/mil)	X^2	P-value
CLEC	1 207 879	606	501.705 88	3.892 (df=1)	0.049
Sub-BAWEC	1 450 308	651	448.870 17		

Note. Raw Freq.=Raw Frequency; Stndd.Freq.(/mil) =Standard Frequency per million texts.

Table 4. Comparison of *although* and *though* frequencies in five sub-CLEC Corpora

Corpus	ST2	ST3	ST4	ST5	ST6	CLEC
Total Freq. (*although & though*)	267	99	98	193	285	942
although Freq.	29 (10.86%)	31 (31.31%)	38 (38.78%)	94 (48.70%)	144 (50.53%)	336 (35.67%)
though Freq.	238 (89.14%)	68 (68.69%)	60 (61.22%)	99 (51.30%)	141 (49.47%)	606 (64.33%)

As shown in Table 3, significant difference exists in general frequencies of *though* in CLEC and sub-BAWEC (P-value=0.049<0.05). Per million texts, more cases of *though* are employed in written texts of CLEC than those of sub-BAWEC. While the general *although* and *though* proportion within sub-BAWEC is 2:1(with 1353 cases of *although* and 651 cases of *though*), it shows a completely opposite trend in CLEC. In concessive clauses where *although*/*though* should be used, there are 606 cases of *though* found in CLEC, almost double the amount of *although* (336). Different from the *although* underuse tendency, Chinese EFL learners demonstrate an overuse tendency of *though* in their writing. Dramatic overuse of *though* is detected in high school students' writing from ST2 as demonstrated in Table 4 where *though* percentage (89.14%) is nearly nine times higher than *although* percentage (10.86%). Significantly high employment of *though* is also found in ST3 and ST4 in which more than 60% of *although*/*though* concessive clauses are introduced by *though*. Employment of *although* and *though* gains a better balance in ST5 but *though* is still overused (around 1:1 in proportion to *although*) compared to the *although*/*though* proportion in sub-BAWEC. ST6 is the only sub-corpus of CLEC where there are slightly more applications of *although* than *though*.

As mentioned, inappropriate selection of wording in written genre can be a factor that leads to the underuse tendency of *although*. Induced interference can account for Chinese EFL learners' general incapability of applying proper wording according to different text genres. It is important for learners to acquire the ability to select proper words for spoken texts and written texts separately. Informal words are mostly used in spoken texts while in formal writing, formal wording should be employed. However, based on the author's knowledge and experience of EFL learning and teaching in China and personal communication with some EFL teachers, in many Chinese English classrooms, students are mainly required to understand the meanings and remember the spellings of the words while the importance of proper wording in different genres is usually less emphasized by language instructors. Hence many Chinese EFL learners may fail to construct a sense of informal and formal words distinction. Specifically, in the case of *although* employment, as Chinese EFL learners, especially learners of relatively low proficiency levels, are more familiar and confident with the informal *though* for it is taught earlier than *although* and with no recognition that *although* is better preferred in written texts, they instinctively tend to use *though* rather than *although* in their writing. Advanced learners especially English majors obtain better control of *although* and *though* employment because they receive highly intensive English language drilling and use English in studies and research as required in the English major academic objectives, and therefore may gain more knowledge on genres distinction and wording selection. As a result, even though an underuse status of *although* still exists in their written output compared to the output of native English speakers, the amount of *although* employment increases significantly.

5.2 Initial and Medial Placement of although

The *although* clauses can be uttered in two forms, namely 1) *Although* P, Q and 2) Q *although* P. The two forms are usually used interchangeably by either native speakers or English learners regardless of their semantic and pragmatic differences that discussed academically by many linguists. Nonetheless, an investigation of Chinese EFL learners' preference of *although* placement selection is beneficial for the analysis of Chinese EFL learners' IL features on *although* employment.

5.2.1 *although* Placement in CLEC and Sub-BAWEC

In the present study, initially placed *although* cases and medially placed *although* cases are spotted primarily by concordance action with "case sensitive" function in the AntConc program, which recognizes *although* and *Although* as two different words. Meanwhile, manual inspection is undertaken to ensure mis-categorized cases due to capitalization errors, juxtaposition, nesting clause complex and other factors are corrected. Table 5 illustrates the frequencies of the initial placement and the medial placement of *although* respectively in CLEC

and sub-BAWEC. 7 cases from CLEC are excluded because *although* is misused in these cases and they are not qualified for the present analysis.

Table 5. General frequencies of *although* placement in CLEC and Sub-BAWEC

	CLEC	Sub-BAWEC	X^2	P-value
Initial *although* Freq.(Pct.)	265 (80.55%)	1041 (76.94%)	1.984	0.159
Medial *although* Freq.(Pct.)	64 (19.45%)	312 (23.96%)	(df=1)	
although Total Freq.(Pct.)	329 (100%)	1353 (100%)		

As reflected by Table 5, about 80% of *although* applications in CLEC is in the form of "*Although* P, Q" and only about 20% of *although* cases are in the form of "Q, *although* P". This largely corresponds to the *although* placement status in Sub-BAWEC in which about 77% of *although* is placed initially in clauses. The chi-square test result manifests that no significant difference is observed on the *although* placement preference from CLEC and sub-BAWEC (P-value = 0.159 > 0.05).

5.2.2 Chinese EFL Learners' *although* Placement Preference

As shown from the quantitative result, there is no significant difference between native English speakers and Chinese EFL learners on their preference of choosing *although* clause forms. The result indicates that Chinese EFL learners possess a relatively high level of nativeness in sentence forms when producing *although* clauses of which they prefer to produce the "*Although* P, Q" form than the "Q *although* P" form. This shares similarity with the native speakers. L1 positive transfer can be a possible explanation for the above finding.

Carbonell-Olivares (2009) confirms the hypothesis in her corpus-based empirical study of meaning and function of *although* that the most cognitively natural ordering of *although* concessive clauses is "the prototypical ordering of subordinate clauses preceding the main clauses", which is the most frequent ordering of *although* clauses for native English speakers. *Although* is usually placed in the initial thematic position to present known/given information that is less prominent or as background information even though authors might sometimes arrange *although* to introduce new information in the medial rhematic position in order to highlight his own claim or opinion. Carbonell-Olivares' conclusion is corresponded with the *although* initial and medial placement observation results in sub-BAWEC.

As for Mandarin Chinese, conjunctions commonly function in pairs. A subordinating conjunction often works with another conjunction to form a clause complex. Some examples include *suiran…danshi…* (literal translation: although…but…), *jinguan…haishi…* (literal translation: in spite of…still…), and *jishi…ye…* (literal translation: even if…still…). It is highly correspondent to English that the subordinating conjunction is placed most cognitively naturally in the initial thematic position to present given/known background or conditional information such as *Suiran kaoshi zhende hen nan, danshi suoyou ren dou hege le* (*Although the exam was really hard, everyone passed*). In some occasions, the subordinating conjunction might be placed in the rhematic position to highlight the author's specific emphases or opinions, such as *Suoyou ren dou hege le, suiran <u>kaoshi zhende hen nan</u>* (*Everyone passed although <u>the exam was really hard</u>*).

Placement preferences of subordinating conjunctions like *although* in English and *suiran* in Mandarin Chinese are cognitively similar and correspondent. Therefore, it is natural for Chinese EFL learners to acquire the same *although* placement preference as that when they produce subordinate clauses in Mandarin Chinese, and thereby approach English nativeness when using *although* clauses in their writing. However, although the finding indicates that Chinese EFL learners can generally master the employment of initial and medial *although*, the slightly higher frequency in *although* initial placement in CLEC than that in sub-BAWEC might imply that Chinese EFL learners are still not fully efficient and confident when producing *although* clause in which *although* is placed medially. Therefore, the uncomfortableness and hesitation in producing "Q *although* P" clauses might also, to some extent, account for the initial placement preference of *although* in Chinese EFL learners' *although* clauses output.

5.3 Variety Ranges of "although" Clause Constructions

As reviewed in 2.1, *although* can introduce non-finite concessive clauses apart from finite clauses, and *although* can be followed by components like noun groups, adjective groups and adjuncts. In other words, variety exists in *although* clause constructions. Moreover, the ability to construct various types of *although* clauses can be a

representation of learners' proficiency level of *although* clause acquisition.

5.3.1 Variety Ranges of *although* Clauses in CLEC and Sub-BAWEC

In order to observe the variety ranges of *although* clauses produced by Chinese EFL learners and conduct CIA between NL and IL texts of *although* clauses, *although* clauses data from both CLEC and sub-BAWEC are extracted and manually categorized into five clause types, namely 1) *although* + finite clause; 2) *although* + non-finite clause; 3) *although* + noun group; 4) *although* + adjective group; and 5) *although* + adjuncts. Frequencies and percentage of each clause type are calculated and chi-square test is conducted to examine whether significant difference exists in *although* clauses variety ranges between writing of Chinese EFL learners and that of native English speakers. The figures and results are presented in Table 6. All the *although* clauses included in the two corpora are involved in the calculation regardless of learner errors.

Table 6. Frequencies of five types of *although* clauses in CLEC and Sub-BAWEC

	al.+ Finite	*al.*+ non-Finite	*al.*+ NG	*al.*+ Adj.G	*al.*+ Adjuncts	Total	X^2	P-value
CLEC	328	3	3	2	0	336		
	(97.6%)	(0.9%)	(0.9%)	(0.6%)	(0%)	(100%)	33.291	
Sub-BAWEC	1178	68	14	72	21	1353	(df=4)	0.000
	(87.1%)	(5.0%)	(1.0%)	(5.3%)	(1.6%)	(100%)		

As demonstrated in Table 6, "*although* + finite clause" is the absolute dominant *although* clause type in both CLEC and sub-BAWEC. However, native English speakers are capable of producing *although* clauses with a wide variety range while Chinese EFL learners tend to construct almost solely "*although* + finite" clauses. "*although* + non-finite clause" and "*although* + adjective group" clauses hold relatively considerable proportions—around 5% respectively—in native English speakers' *although* clause constructions compared to those of Chinese EFL learners with only 0.9% for "*although* + non-finite clause" and 0.6% for "*although* + adjective group". A small amount of "*although* + noun group" and "*although* + adjuncts" clauses can be seen in native English speakers' output while they almost do not exist in Chinese EFL learners' output (considering that all three "*although* + noun group" clauses in CLEC are erroneous). The result of the chi-square test also indicates that a very highly significant difference on *although* clauses variety exists between written output of Chinese EFL learners and that of native English speakers (P-value = 0.000 < 0.001).

5.3.2 Chinese EFL Learners' Mono-structural Tendency in *although* Clause Constructions

As demonstrated in the above quantitative analysis, Chinese EFL learners manifest a mono-structural tendency when constructing *although* clauses in their writing. The "*although* + finite clause" type of *although* clauses, which is the most fundamental and dominant type of *although* clauses, is almost the only clause type that Chinese EFL learners produce. The finding indicates that the proficiency level of Chinese EFL learners' *although* acquisition is not sufficiently advanced for them to produce *although* clauses with syntactical flexibility.

Interlingual difference may be one of the factors that contribute to Chinese EFL learners' mono-structural tendency in *although* clause constructions. The language system of English is synthetic in linguistic typology which means English is a language constructed with a high morpheme-per-word ratio and relies on inflection to express different grammatical categories including numbers, tenses, mood, voice, and so on. However, The Mandarin Chinese language system is analytic and does not rely on lexical inflection to express grammatical relations. Definition and classification of finite verb and non-finite verb which is very important in English does not exist in the Mandarin Chinese system. Therefore, it is difficult for Chinese EFL learners to recognize, distinguish and apply finite and non-finite verbs to form finite or non-finite clauses flexibly. Nevertheless, as finite clauses are frequently drilled throughout the process of English learning, Chinese EFL learners may feel relatively more certain and confident to produce finite clauses. Specifically, in the present study, as Chinese EFL learners are not familiar with both *although* and non-finite clauses, it is reasonable that very few cases of "*although* + non-finite clauses" exist in CLEC.

In addition, if the "*although* + adjective group", "*although* + noun group" and "*although* + adjuncts" clause types are further analyzed, finding will be that all the above clause types can be viewed as the varied or abbreviated categories of the basic "*although* + finite clause" type because the adjective group, noun group and

the adjuncts are all in fact components of a complete finite clause. For instance, among the following *although* clauses, 2), 4), 6) can be modified from 1), 3), 5).

1) *Although Tom was a bit short of breath, he did not give up.*

2) *Although a bit short of breath, Tom did not give up.*

3) *This is a highly democratic country although it is a small one.*

4) *This is a highly democratic country although a small one.*

5) *He managed to complete the task although he did it with tremendous pain.*

6) *He managed to complete the task although with tremendous pain.*

The variety of *although* clause types, in other words, is a representation of syntactical flexibility of English. Compared to English, Mandarin Chinese is less flexible in syntax structure. Clause components cannot be casually omitted and hence the variety range of clauses is narrower. Therefore, when Chinese EFL learners attempt to produce *although* clauses, they may have more difficulties in producing other clause types apart from the basic *although* + finite clause type for they need to adjust themselves to the syntactical flexibility of the English language. This can explain the scarceness of "*although* + adjective group", "*although* + noun group", and "*although* + adjuncts" clause types in CLEC.

To conclude, it is possible that it is the typological and syntactical difference between the English and Mandarin Chinese language systems that results in the mono-structural tendency in Chinese EFL learners' *although* clause constructions.

6. Conclusions

Section 5 explores both quantitatively and qualitatively the characteristics of Chinese EFL learners' *although* employment in written texts from three aspects, and factors that attribute to form these characteristics are discussed. Generally, Chinese EFL learners underuse *although* in their writing and the *although* clauses they produce are highly mono-structural. The manifestation of this status is probably caused by the difference between English and Mandarin Chinese language systems and learners' lack of proper instruction and practice on certain language application skills such as proper wording selection and genres distinction. Despite the above non-nativeness, Chinese EFL learners possess similar *although* placement preference as native English speakers which mainly owe to the similar cognitive process when producing subordinating clauses in English and Mandarin Chinese.

The findings of the above analysis shed light on the improvement of pedagogical practice in Chinese English classrooms. Some pedagogical implications are presented as follows.

First, language instructors should pay more attention to the explanation of language use in different text genres. At present, grammar drilling is the most important subject arranged in most language classrooms and course schedules. However, it is as important for learners to acquire the knowledge of how to use the language appropriately. The proper application of *although* and *though* is only one of the examples among numerous English wording issues in English language use. An increase of attention paid to language use instruction will certainly help learners approach nativeness in L2 production.

Second, when designing and compiling course materials, it is significant for educators and textbook editors to consider learners' specific cognitive style influenced by their NL. Non-nativeness and errors may occur if cognitive patterns on specific language subjects or grammatical process are different in L1 and L2. It is also important to understand the difficulty levels of language items and select proper language items for learning in accordance to learners' proficiency levels to ensure that learners of certain levels are capable of absorbing and applying the new knowledge thoroughly.

The present study only discusses Chinese EFL learners' *although* employment based on their written output and only three selected aspects of *although* employment features are explored. *although* employment in spoken texts should also be analyzed meanwhile more aspects of *although* application should be investigated if a complete and comprehensive understanding of how Chinese EFL learners employ *although* in their L2 output is to be achieved.

References

Aarts, J., De Haan, P., & Oostdijk, N. (Eds.). (1993). *English language corpora: Design, analysis and exploitation*. Amsterdam & Atlanta: Rodopi.

Aijmer, K., Altenberg, B., & Johansson, M. (Eds.). (1996). *Languages in contrast: Papers from a symposium on*

text-based cross-linguistic studies. Lund: Lund University Press.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999/2000). *Longman grammar of spoken and written English*. Beijing: Foreign Language Teaching and Research Press.

Brown, H. (1994/2001). *Principles of language learning and teaching* (3rd ed.). Beijing: Foreign Language Teaching and Research Press.

Carbonell-Olivares, M. (2009). A corpus-based analysis of the meaning and function of *although*. *International Journal of English Studies*, *9*(3), 191-208.

Collins COBUILD English Grammar. (1990). London: William Collins Sons & Co. Ltd.

Corder, S. (1971). Idiosyncratic dialects and error analysis. In J. Richards (Ed.). (1974). *Error analysis: Perspectives on second language acquisition* (pp. 158-171). London: Longman Group Limited.

Corder, S. (1978). Language-learner language. In J. Richards (Ed.). (1978). *Understanding second & foreign language learning: Issue & approaches* (pp. 71-93). Rowley, MA: Newbury House Publishers.

Granger, S. (1993). The international corpus of learner English. In J. Aarts, P. De Haan & N. Oostdijk (Eds.). (1993). *English language corpora: Design, analysis and exploitation*. (pp. 57-71). Amsterdam & Atlanta: Rodopi.

Granger, S. (1994). The learner corpus: A revolution in applied linguistics. *English Today*, *10*(3), 25-29. https://doi.org/10.1017/S0266078400007665

Granger, S. (1996). From CA to CIA and back: An integrated approach to computerized bilingual and larner corpora. In K. Ajjmer, B. Altenberg, & M. Johansson (Eds.). (1996). *Languages in contrast: Papers from a symposium on text-based cross-linguistic studies*. (pp. 37-51). Lund: Lund University Press.

Granger, S. (2003). The international corpus of learner English: A new resource for foreign language learning and teaching and second language acquisition research. *TESOL Quarterly*, *37*(3), 538-546. https://doi.org/10.2307/3588404

Granger, S., & Petch-Tyson, S. (1996). Connector usage in the English essay writing of native and non-native speakers of English. *World Englishes*, *15*, 19-29. https://doi.org/10.1111/j.1467-971X.1996.tb00089.x

Gui, S. C., & Yang, H. Z. (2003). *Chinese learner English corpus*. Shanghai: Shanghai Foreign Language Education Press.

Iten, C. (2000). Although revisited. *UCL Working Papers in Linguistics*, *12*, 65-95.

Mesthrie, R. (Ed.). (2001). *Concise encyclopedia of sociolinguistics*. Oxford: Elsevier Science.

Nemser, W. (1971). Approximative systems of foreign language learners. In J. Richards (Ed.). (1974). *Error analysis: Perspectives on second language acquisition* (pp. 55-63). London: Longman Group Limited.

Paquot, M. (2010). *Academic vocabulary in learner writing: from extraction to analysis*. London: Continuum International Publishing Group.

Pery-Woodley, M. (1990). Contrasting discourses: Contrastive analysis and a discourse approach to writing. *Language Teaching*, *23*, 143-151. https://doi.org/10.1017/S0261444800005772

Quirk, R., Greenbaum, S., Leech, G. & Svartvik, J. (1985). *A comprehensive grammar of the English language*. London: Longman Group Limited.

Richards, J. (Ed.). (1974). *Error analysis: Perspectives on second language acquisition*. London: Longman Group Limited.

Richards, J. (Ed.). (1978). *Understanding second & foreign language learning: Issues & approaches*. Rowley, MA: Newbury House Publishers.

Selinker, L. (1972) Interlanguage. In J. Richards (Ed.). (1974). E*rror analysis: Perspectives on second language acquisition* (pp. 31-54). London: Longman Group Limited.

Spycher, P. (2007). Academic writing of adolescent English learners: Learning to use *although*. *Journal of Second Language Writing*, *16*, 238-254. https://doi.org/10.1016/j.jslw.2007.07.001

Tarone, E. (2001) Interlanguage. In R. Mesthrie (Ed.). (2001). *Concise encyclopedia of sociolinguistics*. (pp. 475-481). Oxford: Elsevier Science.

Web Pages

Net. 1. Middle School English Vocabulary List (People's Education Press Version): 8th Grade-I. Retrieved May 22, 2017 from http://www.xuexila.com/xuexifangfa/banianjiyingyu/2938178.html

Net. 2. Oxford Middle School English Vocabulary List: Grade 9A. Retrieved May 22, 2017, from https://wenku.baidu.com/view/922f734ed0d233d4b04e6987.html

Net. 3. Fifth Grade English Grammar: Inverted Sentences with *as* and *though*. Retrieved June 12, 2017 from http://www.jiaoxue51.com/55768/578432080/161851.html

Note

Note 1. In linguistics studies, the statistical test of significance is frequently used to investigate two or more sets of comparable language data. Conventionally, observed differences with a $P \leq 0.05$ are "significant", with a $P \leq 0.01$ are "highly significant", and with a $P \leq 0.001$ are "very highly significant". The present study also adopts this convention.

Copyrights