

Investigating the English Proficiency of Learners: A Corpus-Based Study of Contrastive Discourse Markers in China*

Guo-Ping Yang¹, Yin Chen²

¹School of Foreign Languages, Beihang University, Beijing, China

²School of International Education and Exchange, Changzhou University, Changzhou, China

Email: gpyang01@163.com

Received 13 November 2014; accepted 19 June 2015; published 24 June 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Discourse markers signal the relationship between the neighboring sentences. The present comparative study investigates the use of Contrastive Discourse Markers between Chinese English learners and native speakers based on corpus data. Special attention is allocated to *but* and *however*, since these two small words are the most popular discourse markers in Chinese English Learners and Native Speakers. Quantitative and qualitative analyses indicate that both groups prefer to employ discourse markers like *but*, *although*, *(even) though*, *however*, etc. when signaling a contrastive relationship between S1 and S2, though they have different priorities in different contextual situation; besides, *but* is overused significantly by Chinese English Learners to signal a contrastive relationship rather than add further information in context; lastly, Chinese English Learner usually employs *however* at the beginning of the sentence, while native speakers put it both at the beginning or middle of the sentences, which both signal the relationship between topics and messages. The findings also suggest that more detailed instructions should be delivered on the procedural meanings and syntactical positions of Contrastive Discourse Markers used in context.

Keywords

Contrastive Discourse Marker, Language Proficiency, Preference, Corpus

1. Introduction

Discourse markers have prevailed in our language expression, which undertake the function of promoting the

*This paper is part of Project of Philosophy and Social Science Research in Colleges and Universities in Jiangsu Province: An Application Study of Corpus-driven Model in the Classroom of Foreign Language Teaching (2013SJB740003).

How to cite this paper: Yang, G.-P., & Chen, Y. (2015). Investigating the English Proficiency of Learners: A Corpus-Based Study of Contrastive Discourse Markers in China. *Open Journal of Modern Linguistics*, 5, 281-290.

<http://dx.doi.org/10.4236/ojml.2015.53025>

expression of communicative intent, and meanwhile facilitate the message interpretation of receivers. In this sense, discourse markers are regarded as a kind of metalanguage which could reflect speaker's metapragmatic awareness (Wu & Yu, 2003) and major difference between the second language learners and native speakers has to do with the frequency of individual markers (Aijmer, 2004). It has already also been proved that the considerable underuse in the non-native group of "small words" (especially among the less native learners) was correlated with their lack of fluency (Hasselgren, 2002). As a matter of fact, a lot of existing researches on the second language acquisition have made kinds of comparisons between native and non-native speaker, and one of the most striking discoveries is that if an second language speaker wants to sound more like a native speaker, one way is to adopt the "conventional expressions" (e.g. Discourse Markers) used by native speakers in the native community (Liao, 2009).

For this reason, a contrastive study between interlanguage and target language is of great significance to the acquisition of English learners in China, due to the pervasive inefficiency of English learning that a considerable gap exists in language proficiency between Chinese English learners and native speakers though they have made great efforts on language learning when they are in school. Furthermore, the findings of this contrastive study could also offer some useful pedagogical implications to language teachers and meanwhile, shed light on the essence of language learning and teaching so as to improve the efficiency of teaching process as well as the teaching effect in classroom. To sum up, the comparison involved in this study could reveal the significant differences between the interlanguage and target language and finally facilitate the understanding of language macroscopically.

Hot topic as it is, it is really difficult to work out a standard definition of discourse marker, or it is even uncertain whether we should call those small words discourse makers, which functions as a marker in discourse, as more than 27 definitions could be found in piles of documents, such as preface (Stubbs, 1983), cue word (Rouchota, 1996), discourse connective (Blakemore, 1987, 1992), conversational routines (Aijmer, 1996), pragmatic formatives/markers (Fraser, 1985, 1996), etc. Despite the fact that the connotation and denotation of those terminologies vary or overlap in degree, due to the distinct research approaches and purposes, more and more scholars have accepted the name of discourse marker proposed by Schiffrin, who define discourse markers as sequentially dependent elements which bracket units of talk (Schiffrin, 1987). It is generally accepted that discourse marker is a category of function words, which includes members of a number of different word classes, including adverbs, connectors, parenthetical expression, as well as particles (Risselada & Spooren, 1998), which shares a common feature that they do not convey propositional meaning but fulfill more procedural functions during the process of meaning expression and interpretation.

1.1. Theoretical Framework of Discourse Marker

Despite the fact that there is no consensus on the definition of discourse marker, the current studies on this topic mainly fall into three categories, namely, Coherence-based Framework (Schiffrin, 1987), Relevance-based Framework (Blakemore, 1987, 1992, 2002), as well as Syntactic/Pragmatic-based Framework (Fraser, 1985, 1996, 1999, 2006). Schiffrin's study originates from the encoding and decoding of message and maintains that discourse markers could facilitate the interpretation of coherence relations between particular units and other surrounding units or communicative situations and they function as contextual coordinators which "indexes an utterance to the local contexts in which utterances are produced and in which they are to be interpreted" (Schiffrin, 1987: p. 326). On the other hand, Blakemore and Fraser have established their studies based on relevance theory and syntactic pragmatic function respectively. The former believes that the major role of discourse markers is to constrain the interpretation of two utterances in context, which could alternatively derive a contextual implication, strengthen an existing assumption, or contradict an existing assumption (Blakemore, 1992). Originating from the speech act theory, Fraser proposes that utterance is a reflection of the speakers' attitude toward communication, and an utterance is usually composed of a proposition and some lexical words which are used to signal the speaker's communicative intent. As the concept of lexical word narrowed down, the terminology is also narrowed down through the whole research process from Pragmatic Formative (Fraser, 1985) to Pragmatic Marker (Fraser, 1996), and finally discourse marker (Fraser, 1999, 2006). In spite of these different understandings to discourse marker, none of the three scholars listed above deny the importance as well as the function of discourse marker in meaning expression and interpretation, which lay a foundation to the present study.

1.2. Working Definition of Discourse Marker in Present Study

Since this study focuses on the signaling function of contrast between different propositions, it adopts the definition of Fraser (2006) as a working definition, namely, *for a sequence of discourse segments S1-S2, each of which encodes a complete message, a lexical expression (LE doesn't contain any propositional meaning) functions as a discourse marker if, when it occurs in S2-initial position (S1-LE+S2), LE signals that a semantic relationship holds between S2 and S1 which is one of: a. elaboration; b. contrast; c. inference; d. temporality* (Fraser, 2006). Corresponding to the four different semantic relationships between S1 and S2, Fraser (2006) classified discourse markers into four distinct groups, including *Elaborative Discourse Markers*, *Contrastive Discourse Markers*, *Inferential Discourse Markers*, as well as *Temporal Discourse Markers*. This working definition reveals the true nature of discourse marker, the feature of non-proposition, and its signaling functions between two propositions, which is the main concern of present study.

Despite the fact that a significant number of previous efforts have been drawn to the theoretical discussion on the essence of discourse marker during the past decades, such as Quirk (1985), Redeker (1990), Traugott (1995), Schourup (1999), Fraser (2006), etc., the descriptive studies on discourse markers are still comparatively limited. This paper will focus on the use of Contrastive Discourse Markers (hereafter CDM for short), a sub-class of discourse marker between Chinese English Speakers and Native English Speakers based on Corpora. Considering the uncertainty of contrast in literature, it is unsurprising that no agreement was reached before on the member of CDMs. Since the theoretical discussion of CDM is not involved in present study, this paper will only concentrate on the discourse markers that generally fall into the concept of *Contrastive*. Part of the frequently-used Contrastive Discourse Marker based on our intuition listed below are the major concern in present study just because it is really a useful and ordinary method in daily utterances, debates, as well as new message introduction, since CDM signals that the explicit interpretation of S2 contrasts with an interpretation of S1. The contrast included here is either between the two literally expressed propositions or between implications or between an implication and a literally expressed proposition (Feng, 2008).

but, however, (even) though, although, on the other hand, on the contrary, whereas, nonetheless, even so, alternatively, conversely, notwithstanding

2. Research Questions

The present study presupposes that Chinese English learners and native speakers differ in speaking styles with different preferences to discourse markers when signaling contrastive relationships. Compared with native speakers, Chinese English learners may overuse or underuse certain CDMs in expression, which makes them sound less native than native speakers even though they are proficient English users, since discourse markers are difficult for second language learners to acquire without any exposure to the target language environment. Therefore, the purpose of this study was to investigate the preference or priority to different CDMs in use between two different ethnic groups, and specifically, the following questions were addressed:

- Whether particular CDMs will take priority in the choice of different group speakers?
- Do Chinese English learners overuse or underuse one or two CDMs compared to native speakers?
- Do Chinese English learners use CDMs for the same purpose as native speakers? Is there any difference on the procedural functions of CDMs in context when a same CDM is employed in the use of two different groups?
- Do they use CDMs at the same syntactical position in the sentence? Is there any preference?

3. Research Methodology

The material employed in this study come from CLEC (Chinese Learner English Corpus) and FLOB (Freiburg-LOB Corpus of British English), which both covers more than a million words. The language material is processed through Wordsmith 4.0 through the function of *Wordlist* and *Concordance*.

The frequency of each CDM listed above was retrieved respectively in the two corpora to find which CDMs were comparatively more significant than others in the data in order to study the preference of different ethnic groups in language use. Two typical discourse markers were selected from the sentences retrieved in the two corpora as examples to analyze their main functions as well as their syntactical positions in the sentences of corpora, so that a clear comparison could be made on the usage of these two CDMs between Chinese English Learners and native speakers.

4. Data and Findings

The concern of the whole comparative study on CDM in this paper could be summarized into one question—preference, including the preference to the choice of CDMs in the surface level when signaling the contrastive relationship between two neighboring sentences, and the preference to the position of a CDM in the surface level when it is employed in the text, as well as the preference to the procedural function of a certain marker in the deep level in utterance. In other words, this study was committed to the following question: to a certain discourse slot when indicating a contrastive relationship, which CDM should be chosen, for what purpose and where it should be put.

4.1. Preference Analysis

Twelve CDMs listed above in all are examined during the study. The frequency of each in CLEC and FLOB are recorded and the log-likelihood are calculated during the process to test the significance of each CDM in the study, the result of which are displayed together in the **Table 1** so that we can detect the preference of each ethnic group at a glance.

The CDMs are arranged in a decreasing order of frequency based on the data collected from Chinese English learners, namely, *but*, *however*, *(even)though*, *although*, *on the other hand*, *on the contrary*, *whereas*, *nonetheless*, *even so*, *alternatively*, *conversely*, *notwithstanding*, and *contrariwise*, etc., which form a sharp contrast with its corresponding data from FLOB. The statistics have also offered another powerful circumstantial evidence to the presupposition we have mentioned above that both the Chinese English learners and native speakers differ in preference among tremendous CDMs when signaling contrastive relationships.

As to the frequency of each CDM in different group of speakers, the study reveals a similar tendency of preference on the whole in both groups when signaling contrastive relationships and the option of CDM is generally concentrated on several choices, though the frequency of each varies, if those CDMs whose frequency is less than 10 times in the corpus statistics are ignored. As a matter of fact, the CDMs tested in the study could be classed into three groups based on their frequency in both corpora: *but* comes first, while *however*, *although*, *(even) though* come as the second group, and *on the other hand*, *on the contrary* and *whereas* as the third group (see **Table 1**). Besides, the choice of CDMs to Chinese English learners are comparatively limited, as *even so* and *nonetheless* lose their traces in CLEC, even though they are also not very popular among native speakers. For this reason, the distribution of each CDM in FLOB is more reasonable than CLEC and the choice of native speakers when signaling contrastive relationship is much more diversified than Chinese English Learners though the frequency of each word varies.

Table 1. Frequency of CDMs.

CDM	Freq. in CLEC	Freq. in FLOB	Log-likelihood	Sig. ($p < 0.05$)
But	3327	1548	847.01	0.000 *** +
However	592	566	5.77	0.016 * +
(Even) Though	333	540	31.65	0.000 *** -
Although	428	597	14.10	0.000 *** -
On the other hand	72	54	4.61	0.032 * +
On the contrary	31	7	18.77	0.000 *** +
Whereas	17	15	0.39	0.531 +
Conversely	4	11	2.76	0.097 -
Notwithstanding	2	9	4.17	0.041 * -
Alternatively	1	13	11.08	0.001 *** -
Nonetheless	0	9	#NUM!	#NUM! ### -
Even so	0	16	#NUM!	#NUM! ### -

* stands for the significance of statistics; # means null or the statistics here is invalid or meaningless

Eight CDMs in twelve in the study have significantly differences between two corpora. Particularly, in terms of log-likelihood ratio, Chinese English learners prefer to *but*, *however*, *on the other hand*, and *on the contrary*, while native speakers are more inclined to use *(even) though*, *although*, *notwithstanding* as well as *alternatively*. In addition, considering capacity proportion, the CDM *however* is much more popular among native speakers, though the absolute figure of *however* in CLEC outweighed in FLOB. Despite the differences, these CDMs constitute a regular corpus of signaling contrastive relationships between S1 and S2, and different speakers could choose what they need for a certain discourse slot based on the different education they received or the learning environment.

Furthermore, if we reorder the CDMs above based on the frequency, the significant differences between these two groups will stand out (See **Table 2**) that both groups would give first priority to *but*, *although*, *(even) though*, *however*, etc. when signaling relationships between S1 and S2. Actually, the four CDMs outweigh all the other CDMs in the two different groups of English speakers. This conclusion coincides with the division of cancellative discourse markers in [Bell \(1998\)](#) who classified *but*, *however*, and *though* as primary core cancellatives. Besides, the marker *but* is overused by Chinese English Learners, which makes a sharp contrast with the use of the other three, due to a negative transfer of Chinese, since it is really a first choice to most Chinese speaker when they want to deliver some negative ideas in their mother tongue. It has also been proved by [Wang & Zhu \(2005\)](#), who focuses on the features of discourse marker in oral English of Chinese Speaker. Comparatively speaking, though the native speakers' choices are diversified, including *but*, *however*, *(even) though*, *although*, etc. the word *but* in FLOB is not employed as frequently as it in CLEC, and the frequency of the latter three words listed when added together equals to that of *but* in native speakers.

Last but not least, the CDMs involved in this study mainly focus on the semantic meaning *CONTRAST*, including a contrast of the explicit message of S2 with the explicit or implicit message of S1, or an implication between the two sentences that the message conveyed by S1 is false, while S2 is correct. In spite of similar functions with the core meaning of contrast, the procedural meanings of those CDMs vary in sentences, which exert restrictions on the specific relationships between S1 and S2. It is generally accepted that those CDMs with least restrictions on relationships would usually be recruited in talks unless a particular requirement cannot be fulfilled. For this reason, the frequency of each CDM in corpora (See **Table 2**) could also be taken as a simple

Table 2. Contrast on the Frequency of CDM in Corpora.

Frequency	CLEC	FLOB
3000 - 4000	But	
2000 - 3000		
1000 - 2000		But
500 - 1000	However	Although However Even though
400 - 500	Although	
300 - 400	Even though	
200 - 300		
100 - 200		
50 - 100	On the other hand	On the other hand
0 - 50	On the Contrary Whereas Conversely Notwithstanding Alternatively	Even so Whereas Alternatively Conversely Notwithstanding Alternatively

reflection on the restriction of each CDM. The CDM frequently used is less restrictive than those infrequently used. Take *but* and *however* for example, the popularity of *but* in both corpora results from the least restrictions it imposes than *however* on the relationship between S1 and S2 which is contrasted, as “*but* seems to identify a matter-of-fact denial, while *however* conveys a kind of reluctance” (Fraser & Malamud-Makowski, 1996).

4.2. Procedural Meaning and Syntactic Position: *But* & *However*

Besides the frequency of discourse markers discussed above, the procedural meaning and syntactical position of discourse markers are another two inevitable questions, since they all work together to decide whether the sentences sounds natural or not.

In order to clarify the significant differences between Chinese English learners and native speakers, two typical CDMs *but* and *however* (600 items in all on *but* and *however* both from CLEC and FLOB selected respectively at random) are examined in this study, despite that they both signal a general relationship of contrast. However, the following two cases of *but* and *however* will not be taken into consideration of this study: “*but*” cannot be classified as CDM, when it means “*except*”, thus, the phrases, such as *nothing but...*, *have no choice but...*, *but to do something...*, *but for...*, etc. are excluded from this study; “*however*” cannot be classified as CDM, when it is used before an adjective or adverb to emphasize that the degree or extent of something cannot change a situation or to indicate that the figure you have just mentioned may not be accurate, etc.

Discourse markers evolved from non-discourse marker sources through historical process of grammaticalization which alter their original meaning (Schourup, 1999). The non-defining feature of any discourse marker’s semantic meaning will be eliminated during the process of grammaticalization until to the core meaning—an invariant semantic content of each marker. It is generally accepted that every discourse marker has only one vague core meaning, whose interpretation is in connection with the entire contextual meaning conveyed by an utterance in which a discourse marker appears and the process of meaning expression as well as utterance interpretation will contribute to the core meaning of proposed discourse marker. In this sense, discourse markers are multi-functional with a stable core function or core pragmatic function and those instantiations outside the core to be peripheral (Bell, 1998). As a matter of fact, many existing studies have argued that the general procedural meaning could be implemented in different ways to derive these meaning. For instance, Hussein (2008) summarized four different meaning encoded under the umbrella of *but*, including a denial-of-expectation meaning between the two conjuncts it links (Blakemore, 1987, 2002), a simple contrast of the relation between the two conjuncts (Lakoff, 1971), a correct placement for the assumption given in the first clause (Anscombe & Ducrot, 1977), and a return to the main topic of discourse. The situation goes the same to the study of *however*, though Fraser (1997) and Fraser and Malamud-Makowski (1996) believes *however* differs from *but* slightly.

In this study, four types of procedural meaning of *but* are collected from the samples of corpora (see Table 3), namely, contrasting with what you have just said, adding to what you have just said or something further in a discussion or returning to the subject, making an excuse or apologized for what you are just about to say. Two types of procedural meaning of *however* are collected likewise, namely, contradicting with the former message, contradicting with something said previously.

It is revealed that the first two procedural meaning of *but* take up a dominant role both in CLEC and FLOB,

Table 3. Procedural Meaning of *But* and *However*.

		But (300)		However (300)			
		CLEC	FLOB	CLEC	FLOB		
	Contrasts with former message	50.67%	38.67%	Contrasts with former message	88.67%	85.33%	
Procedural meaning (Introduce something which...)	Adds further information	38.67%	48.67%	Procedural meaning (Introduce something which...)	Contradicts with something said previously	12.00%	14.67%
	Returns to the subject	9.33%	12.00%				
	Make an excuse or apologize	1.33%	2.00%				

however, the importance of each procedural meaning varies slightly in Chinese English Learners and Native Speakers (see **Table 3**). Chinese Learners are accustomed to giving priority to *but* when they intend to contrast the new message of S2 with the former message from S1 (50.67%) (See **Table 4**), on the contrary, native speakers prefer to employ *but* when introducing further information (48.67%) (See **Table 5**). Furthermore, both Chinese learners and native speakers have got similar understanding on the procedural meaning of *however*, as the function of contrasting with former message outweighs the function of contradicts with something said previously in both groups.

As a matter of fact, the correct use of discourse markers involves a comprehensive understanding to the procedural functions in accordance with native speakers. The statistics in this study indicates that a natural language cannot be acquired without a subtle understanding of the core procedural meaning of discourse markers in target language.

Besides procedural meaning of CDM, syntactical position is another important factor that determines whether the language of English learner sounds nature or not. Fraser (1997) has presented a standard syntactical position of discourse marker, namely a declarative sentence, then a Discourse Marker, followed by a second declarative sentence, and summarized three types of syntactical positions of discourse markers in utterance, i.e. [S1, DM + S2], [S1. DM + S2], and [DM + S1, S2]. Since it is not a theoretical discussion on discourse marker, the study here will not involve the cases of empty S1 or, even sometimes, empty S1 and S2 demonstrated in Fraser (2001). As to the syntactical position of *but*, it is revealed that a large majority of both groups prefer to the structure of [S1, DM + S2], while it differs in the case of *however* (see **Table 6** and **Table 7**). The structural distribution of [S1, DM + S2] and [S1. DM, + S2] in the frequency of *however* equals nearly in FLOB though the first structure gets a little advantage, on the contrary, the latter structure has an overwhelming superiority than the former one in Chinese English, which results from a strong negative transfer of Chinese structure *suiran* (although)... *danshi* (but)... and a lack of language flexibility after a long-term disciplined and rigid language training in China. It is not surprising that the structure of [DM + S2, S1] in *but* and the structure of [S1 + S2, DM] in *however* disappear in the samples checked here, because it does not conform to the speaking style of both groups. In general,

Table 4. *but* in CLEC for contrasting.

success of the Three Musketeers lies in many sides, **but** I think the most important one is that him, nearly jokingly. He was not that nervous, **but** was still embarrassed to be caught red-handed. robot to do the job, I must say "congratulations", **but** now I can be no more the crazy noise. also cook many good things for the ghosts to "eat", **but** in fact they eat the food all by themselves! from shops. The Dragon Boat Festival is traditional, **but** it is full of lives. It will be handed down from Tsinghua University have some collective characters, **but** everyone of them is unique. Trust is often thought to be similar with "to believe", **but** in fact it's more than that. that there are all kinds of fake commodities around us, **but** it is the fact. More and more people have is lower than good one, and you know it is not smart, **but** you also buy it, because of its prize. be seen that someone is too eager to do things ahead, **but** they fail in the end. Why does this happen?

Table 5. *but* in FLOB for adding information.

car?" Sally asked suddenly. It was a car, **but** it was going in the wrong direction. Original by the look of it, and so were the banisters, **but** someone had painted them a kind of snot green. I smile at him and say I don't know, **but** it Really is Appalling'." But what was see only in the tabloids' own opinion columns, **but** it's a view I've heard from a number of senior service is going to be cut is the only news I have, **but** it won't affect you rich men in your big cars." with devastating effect. They adored each other, **but** the odd thing was that, as Thomas aged with not how you want to play then it's all right with me. **But** make no mistake, Marie, we belong together, She wondered why more people weren't like him, **but**, there again, it might be difficult if everyone briefly in the reflected lights on the dashboard. "**But** then, neither are a lot of the things I like

Table 6. Syntactical position of *but* and *however*.

		but (300)		however (300)		
		CLEC	FLOB		CLEC	FLOB
	S1, DM + S2	22.00%	26.67%	S1, DM, + S2	8.667%	54.667%
Position	S1, DM + S2	78.00%	73.33%	Position	S1, DM, + S2	91.333%
	DM + S2, S1	0.000%	0.000%	S1 + S2, DM	0.000%	0.000%

Table 7. Syntactic Structure of *but* and *however*.

phone call from the President of America. “**But** you’re not to tell anyone”, I told her, “because I decent music” can put you right off your groove. **But** of course an e-acuteitist door policy is supposed bimbettes were keen on getting in on all this. **But** they were experiencing difficulties in persuading when a person needsome pills for his sick, **but** he bought fake ones. His illness would We may think they are common and indifferent, **but** when we buy something fake or of bad quality in need of a medicine which to heal his sick, **but** he bought a fake medicine, so after he eating the to expect a little romance between the two. **However**, it wasn’t to be. They spent long was certainly a cut above the Perkins family. **However**, Mrs Saunders had recently realised that her puresnow, there are still crime, famine and war. **However**, human beings are making progress in both except trying not to show her underwear. This bit, **however**, was going to be tricky. She bent slowly that by a mere youth. Caution whispered to him, **however**, that the young man was nobly born, Time goes on its way, we only use time, **however**, we don’t create it. we must make it good us

Chinese English learners have a better mastery of the structure of *but* than that of *however*, which makes them sounds more natural.

4.3. Summary of Findings

The comparative study conducted here focuses on the preferences of CDMs in different language groups as well as their language habits on the syntactical position and procedural meaning of CDMs used. Special attention was allocated to *but* and *however*, since these two small words are more popular discourse markers in Chinese English Learners and Native Speakers respectively. The above analysis indicated that both groups similarly showed a strong inclination on the choice of CDMs, which concentrated on *but*, *however*, *although*, (*even*) *though*, despite the different frequent occurrences of each word resulted from the impact of different meta-awareness.

Besides, significant differences were found through the statistics comparisons that the proficiency of Chinese English learners still needs improving. To begin with, in comparison with the diversified selection of CDMs when signaling contrastive relationship, *But* is overused excessively by Chinese English Learners. The quantitative analysis reveals that superordinate terms are employed more frequently by English learners to cover the hyponyms within the same semantic meaning categories, due to a lack of knowledge on the subtle differences of CDMs.

Furthermore, differences on the understanding of pragmatic functions of CDMs between interlanguage and target language are also reflected from the statistical differences on the procedural meanings between two groups, despite the high-frequent occurrence of *but* in CLEC. Considering the fact that the metapragmatic awareness of speakers, which could predict academic achievement as a kind of cognitive ability (Phillipson & Phillipson, 2012) are embodied in the selection of procedural meaning of CDMs, the data differences in comparison reflect a gap on the understanding of CDM which impedes the acquisition of native language.

Last but not least, an appropriate syntactical position of CDM symbolizes a correct acquisition of target languages in form and an observance of language rule will make the interlanguage sound more natural. Chinese English Learner usually employs *however* at the beginning of the sentence to contrast or contradict with former

information, while native speakers put it both at the beginning or middle of the sentences or sometimes at the end of the sentence, even though no sample is selected due to a random sampling, which is used to signal both the relationship between topics and messages.

In short, the performance of language ability is high-positively correlated with the employment of DM, as Wei (2011) indicates that advanced students were generally more active than intermediate students in using markers.

5. Conclusions

Discourse markers are employed in communication to build discourse coherences, constrain the relationships between two neighboring sentences and facilitate the understanding of utterances. Many existing theoretical and empirical investigations on discourse marker have already revealed that discourse markers is one of the language devices used to signal the metapragmatic awareness of speakers that reflects a regulation on awareness of language, so that the utterance changes as the meta-awareness varies. For this reason, a corpus-based cross-linguistic study on CDM could enable us to detect the differences of different language speakers when speaking, foster a correct understanding and use of CDM in target language, and finally improve the second language proficiency of learners.

This study added empirical evidence both to the improvement of foreign language teaching as well as learning. A great significance should be attached to different types of discourse markers, including CDMs during the process of English learning and teaching. However, it should be noted that Chinese English learners is characterized with the features of interlanguage, since several CDMs, *but* in particular, are overused in expression. Thus, comprehensive instructions should be given to Chinese English learners so as to broaden the scope of discourse markers as well as the procedural functions and syntactical position of each discourse marker in target language. Only in this way, the utterances produced could sound more natural to native speakers.

Consequently, despite the fact that empirical study has offered a new insight into discourse markers, it was by no means intended to provide final answers to the questions addressed before, due to the random sampling in the research. In addition, the correlation between procedural meaning and syntactical position of CDMs is also not considered in the study. Therefore, the findings of this tentative empirical study are expected to be supplemented with a large scale of experimental data in the forthcoming days.

Endnote

Here DM refers to Discourse Marker. S1/S2 in this study includes not only declarative sentences but also interrogative sentences as well as imperative sentences as long as they can work together to signal the contrastive relationship.

References

- Aijmer, K. (1996). *Conversational Routines in English: Convention and Creativity*. London, New York: Longman.
- Aijmer, K. (2004). Pragmatic Markers in Spoken Interlanguage. *Nordic Journal of English Studies*, 3, 173-190.
- Anscombe, C., & Ducrot, O. (1977). Deuxmais en Francais. *Lingua*, 43, 23-40.
[http://dx.doi.org/10.1016/0024-3841\(77\)90046-8](http://dx.doi.org/10.1016/0024-3841(77)90046-8)
- Bell, D. M. (1998). Cancellative Discourse Markers: A Core/Periphery Approach. *Pragmatics*, 8, 515-541.
<http://dx.doi.org/10.1075/prag.8.4.03bel>
- Blakemore, D. (1987). *Semantic Constraints on Relevance*. Oxford: Blackwell.
- Blakemore, D. (1992). *Understanding Utterances*. Oxford: Blackwell.
- Blakemore, D. (2002). *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511486456>
- Feng, G. W. (2008). Pragmatic Markers in Chinese. *Journal of Pragmatics*, 40, 1687-1718.
<http://dx.doi.org/10.1016/j.pragma.2008.05.012>
- Fraser, B. (1985). Pragmatic Formatives. In J. Verschueren, & M. Bertuccelli-Papi (Eds.), *The Pragmatic Perspective* (pp. 179-194). Amsterdam: Benjamins
- Fraser, B. (1996). Pragmatic Markers. *Pragmatics*, 6, 167-190. <http://dx.doi.org/10.1075/prag.6.2.03fra>
- Fraser, B. (1997). Contrastive Discourse Marker in English. In Y. Ziv, & A. Jucker (Eds.), *Pragmatics and Beyond: Discourse Markers* (pp. 301-326). Amsterdam: Benjamins.

- Fraser, B. (1999). What Are Discourse Markers? *Journal of Pragmatics*, 31, 931-952. [http://dx.doi.org/10.1016/S0378-2166\(98\)00101-5](http://dx.doi.org/10.1016/S0378-2166(98)00101-5)
- Fraser, B. (2001). The Case of the Empty S1. *Journal of Pragmatics*, 33, 1625-1630. [http://dx.doi.org/10.1016/S0378-2166\(01\)00024-8](http://dx.doi.org/10.1016/S0378-2166(01)00024-8)
- Fraser, B. (2006). Toward a Theory of Discourse Markers. In F. Kerstin (Ed.), *Approaches to Discourse Particles* (pp. 189-204). Amsterdam: Elsevier.
- Fraser, B., & Malamud-Makowski, M. (1996). English and Spanish Contrastive Discourse Markers. *Language Science*, 18, 863-881. [http://dx.doi.org/10.1016/S0388-0001\(96\)00052-6](http://dx.doi.org/10.1016/S0388-0001(96)00052-6)
- Hasselgren, A. (2002). Learner Corpora and Language Testing: Small Words as Markers of Learner Fluency. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer-Learner Corpora, Second Language Acquisition, and Foreign Language Teaching* (pp. 143-173). Philadelphia, PA: John Benjamins.
- Hussein, M. (2008). The Discourse Markers "But" in English and Standard Arabic: One Procedure and Different Implementations. <http://semarch.linguistics.fas.nyu.edu/Archive/TAYOTBkM/But%20as%20a%20General%20Procedure.pdf>
- Lakoff, G. (1971). The Role of Deduction in Grammar. In J. Fillmore, & D. T. Langendoen (Eds.), *Studies in Linguistic Semantics* (pp. 62-70). New York: Holt, Rinehart and Winston.
- Liao, S. (2009). Variation in the Use of Discourse Markers by Chinese Teaching Assistants in the US. *Journal of Pragmatics*, 41, 1313-1328. <http://dx.doi.org/10.1016/j.pragma.2008.09.026>
- Phillipson, S., & Phillipson, S. N. (2012). Children's Cognitive Ability and Their Academic Achievement: The Mediation Effects of Parental Expectations. *Asian Pacific Education Review*, 133, 495-508. <http://dx.doi.org/10.1007/s12564-011-9198-1>
- Quirk, R. (1985). *A Comprehensive Grammar of the English Language*. London and New York: Longman.
- Rouchota, V. (1996). Discourse Connectives: What Do They Link? *UCL Working Papers in Linguistics*, 8, 1-15.
- Redeker, G. (1990). Ideational and Pragmatic Markers of Discourse Structure. *Journal of Pragmatics*, 143, 367-381. [http://dx.doi.org/10.1016/0378-2166\(90\)90095-U](http://dx.doi.org/10.1016/0378-2166(90)90095-U)
- Schiffrin, D. (1987). *Discourse Markers*. New York: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511611841>
- Schourup, L. (1999). Discourse Markers. *Lingua*, 107, 227-265. [http://dx.doi.org/10.1016/S0024-3841\(96\)90026-1](http://dx.doi.org/10.1016/S0024-3841(96)90026-1)
- Stubbs, M. (1983). *Discourse Analysis: The Sociolinguistic Analysis of Natural Language*. Chicago, IL: The University of Chicago Press.
- Traugott, C. (1995). *The Role of the Development of Discourse Makers in a Theory of Grammaticalization*. Resource Document, Paper Presented at ICHL XII. http://www.wata.cc/forums/uploaded/136_1165014660.pdf
- Wang, L. F., & Zhu, J. H. (2005). A Study on the Use of Discourse Markers in Chinese Students' Oral English. *Foreign Languages Research*, 3, 40-48.
- Wei, M. (2011). Investigating the Oral Proficiency of English Learners in China: A Comparative Study of the Use of Pragmatic Markers. *Journal of Pragmatics*, 43, 3455-3472. <http://dx.doi.org/10.1016/j.pragma.2011.07.014>
- Wu, Y. X., & Yu, G. D. (2003). Metapragmatic Analysis of Discourse Marker. *Foreign Languages Education*, 4, 16-19.