LEARN Journal: Language Education and Acquisition Research Network
(ISSN: 2630-0672 (Print) | ISSN: 2672-9431 (Online)
Volume: 15, No: 2, July – December 2022
Language Institute, Thammasat University
https://so04.tci-thaijo.org/index.php/LEARN/index

A Corpus-Based Study on the Use of Spoken Discourse Markers by Thai EFL Learners

Zhaoyi Pan[a,*], Wirote Aroonmanakun[b]

[a] zhaoyipan1989@gmail.com, English as an International Language Program, Graduate School, Chulalongkorn University, Thailand
[b] awirote@chula.ac.th, Department of Linguistics; Corpus Linguistics for Digital Humanities Research Unit, Faculty of Arts, Chulalongkorn University, Thailand
[*] Corresponding author, zhaoyipan1989@gmail.com

APA Citation:
Pan, Z. & Aroonmanakun, W. (2022). A corpus-based study on the use of spoken discourse markers by Thai EFL learners. *LEARN Journal: Language Education and Acquisition Research Network*, *15*(2), 187-213.

Received
03/02/2022

Received in revised form
15/03/2022

Accepted
28/03/2022

Keywords
spoken discourse markers; pragmatic functions; Thai EFL learners; learner corpus

Abstract

This research investigated the use of English spoken discourse markers by Thai EFL learners in English conversation compared to native English speakers from two perspectives: frequency and pragmatic function. A total of 60 learners were involved in the research: 30 Thai B1-level learners and 30 Thai C1-level EFL learners. Spoken data was collected and transcribed into written form to build a learner corpus for analysis. The data analysis indicated underuse by Thai EFL learners of four spoken discourse markers – *so*, *well*, *you know* and *I think*. Moreover, Thai EFL learners we shown to use each spoken discourse marker differently in comparison to native English speakers. On the whole, interpersonal functions were less frequently a factor than textual functions, indicating a larger deficiency in performing interpersonal functions by Thai EFL learners. These results lead to the conclusion that Thai EFL learners lack pragmatic competence in oral communication in terms of performance (usage instance) discrepancy regarding

	spoken discourse markers compared to native English speakers.

Introduction

Discourse markers have been studied in both written form and spoken form. It has been found that, in spoken form, discourse markers reveal substantially more pragmatic functions than those in written form (Aijmer, 2016). They appear in oral communication with high frequency (Aijmer, 2016; Bolden, 2015), and their pragmatic functions likewise play important roles in oral communication (House, 2013).

Hence, a significant aspect of studying spoken discourse markers is examination of their unique pragmatic functions in communication (Aijmer, 2004, 2011). In this light, previous research was aimed primarily at studying the pragmatic functions of spoken discourse markers used by native English speakers.

Recent studies, however, have tended to focus on comparisons between native English speakers and non-native English speakers across different English levels (Buysse, 2012; Diskin, 2017; Fung & Carter, 2007; House, 2013; Polat, 2011). Two major problems have emerged. The first problem pertains to differences in the frequency of spoken discourse marker use between non-native English speakers and native English speakers (Aijmer, 2004). The second problem comprises the different functional uses of discourse markers between native English speakers and non-native English speakers (Polat, 2011).

There exist some studies on written English discourse marker use by Thai EFL learners (Jangarun & Luksaneeyanawin, 2016; Prommas & Sinwongsuwat, 2011; Sitthirak, 2013); however, few studies have been conducted with aim of determining how Thai EFL learners use spoken discourse markers in English conversation (Arya, 2020; Nookam, 2010).

Under these circumstances, this research was undertaken with two main objectives: 1) to compare the significant differences in the use of spoken discourse markers between Thai EFL learners and native English speakers, and, 2) to identify different pragmatic functions of spoken discourse marker use by Thai EFL learners in comparison to native English speakers.

Literature review

The literature review is separated into four sections: a review on the definition of discourse markers, comparative studies on spoken discourse markers, Brinton's (2008) framework of the pragmatic functions of discourse markers, and studies on English discourse markers used by Thai EFL learners. Each review is presented in its turn.

Definition of discourse markers

In order to define discourse markers, many researchers have placed focus on the interpersonal functions of those markers. Traugott and Dasher (2002) defined discourse markers as signals of "an aspect of the speaker's rhetorical stance toward what he or she was saying, or toward the addressee's role in the discourse situation" (p. 152). This definition takes into account two important pieces of information. Firstly, discourse markers are used to show the attitude of a speaker, and, secondly, a speaker reveals his or her attitudes about or purpose for interacting with the hearer. Bazzanella (2006) proposed a definition from this perspective, saying that discourse markers were useful in "locating the utterance in an interpersonal and interactive dimension… and in marking some on-going cognitive processes and attitudes" (p. 456). This idea aligns with the findings of other research (Fung & Carter, 2007; Sakita, 2013).

Subsequently, Brinton (2008) proposed a definition for discourse markers that is used in this research. Brinton (2008) stated that discourse markers are "phonologically short items that have no or little referential meaning but serve pragmatic or procedural purpose" (p.1). That is, discourse markers are short items in terms of phonology. Moreover, they should have pragmatic or procedural meanings in context rather than possess referential meanings.

Comparative Studies on Spoken Discourse Markers

Experimental studies have focused mainly on comparing discourse marker use between native English speakers and non-native

English speakers. Various studies focused on EFL learners with different backgrounds, such as Hong Kong EFL learners (Fung & Carter, 2007), Swedish EFL learners (Aijmer, 2011), Belgian native speakers (Buysse, 2012), Turkish EFL learners (Asik & Cephe, 2013), EFL learners from Ireland (Diskin, 2017), etc. Most studies have focused primarily on how non-native English speakers make use of the pragmatic functions of certain spoken discourse markers as compared to native English speakers (Aijmer, 2004, 2011; Buysse, 2012).

Several studies have scrutinized the frequency list of discourse markers used by non-native English speakers (Aijmer, 2004; Asik & Cephe, 2013; Fung & Carter, 2007). The main purpose in these works was identifying the most-used discourse markers.

Fung and Carter (2007) studied Hong Kong EFL learners to examine whether they performed spoken discourse markers the same way as did native English speakers. It was found that Hong Kong EFL learners used fewer functions of spoken discourse markers than native English speakers, e.g., *well*, *you know*, indicating a lack of pragmatic competence in oral communication.

A frequency list of the most-used discourse markers was compiled in Aijmer (2004). The top markers were *I think*, *well*, *I don't know*. It was found that EFL learners at less advanced levels seldom used these discourse markers. Furthermore, some learners did not understand the importance of these markers in conversation.

Asik and Cephe (2013) explored Turkish EFL learners' use of discourse markers. Discourse markers used with high frequency were *and*, *umm* and *so*. The study found that Turkish EFL learners used fewer discourse markers than native English speakers.

In addition to the above studies, several other studies have investigated specific discourse markers used by non-native English speakers (Aijmer, 2011, 2016; Buysse, 2012; Diskin, 2017; Polat, 2011). Two major, and seemingly contradictory findings emerged from these studies. At least one found that non-native English speakers used spoken discourse markers with much lower frequency in oral communication (Polat, 2011), whereas others found that non-native English speakers used spoken discourse markers with much higher frequency in oral communication (Aijmer, 2011; Buysse, 2012).

Polat (2011) focused on the use of the discourse marker *well* by immigrant second-language learners. The study shows that learners did

not use *well* as a discourse marker at all, indicating that, to some extent, non-native English speakers did not have an awareness of using certain English discourse markers in oral communication.

Diskin (2017) focused on the use of *like* by non-native English speakers in Ireland. It was found that, compared to native English speakers, learners at a low-level stages of English acquisition used *like* with much less frequency.

As revealed by Fung and Carter (2007), Hong Kong EFL learners exhibited restricted use of spoken discourse markers, including *well*, *you know* and others.

Unlike the results above, Aijmer (2011) reported that Swedish learners, in fact, overused *well* compared to native English speakers. This result was in line with Müller (2005), who found that German learners, likewise, overused *well*.

Although Polat (2011) determined that non-native English speakers did not use *well* as a discourse marker at all, two other markers, namely *you know* and *like*, were used far more frequently by non-native English speakers than by native English speakers.

Buysse (2012) studied *so* as a discourse marker in both EFL learners and native English speakers. The result showed that EFL learners used *so* much more frequently than did native English speakers.

The conclusion to be drawn from the previous research is that non-native English speakers use spoken discourse markers differently than do native English speakers, and that they do so in terms of both frequency and function. It can also be ascertained from previous research that non-native English speakers' lack of comprehensive language acquisition will, in many cases, account for differences in spoken discourse marker use. All told, through the observations and experiments of these and other studies, researchers have identified and elaborated on the different circumstances that influence non-native English speakers' pragmatic competence in oral communication.

Brinton's (2008) Framework of the Pragmatic Functions of Discourse Markers

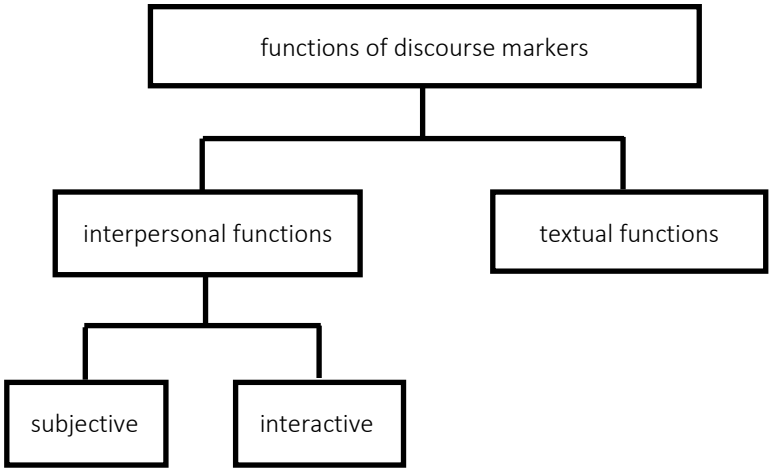
This research adopts Brinton's (2008) framework of the pragmatic functions of spoken discourse markers because it is a complete and concise framework. Brinton developed this framework based on

Schiffrin's five planes of talk (1987) and Relevance Theory (Aijmer, 2011). Brinton further consulted the findings of previous empirical studies in constructing the framework (Bazzanella, 2006; Brinton, 1996).

Discourse marker functions can be separated into two main perspectival domains. The first domain is from the interpersonal perspective, which reveals the relationship between the speaker and the hearer. Brinton (2008) separates interpersonal functions into two groups: subjective functions, relating to the expression of a speaker's attitude, and interactive functions, relating to actions made by a speaker towards the hearer between utterances. The second domain, moreover, forms a textual perspective and highlights the relationship between prior and subsequent utterances. Brinton's framework of functions of discourse markers is presented in Figure 1.

Figure 1

Brinton's framework of functions of discourse markers (2008)



Based on this framework, Brinton (2010) analyzed the discourse marker *I mean*. Examples of both *interpersonal functions* and *textual functions* retrieved from The Freiburg-LOB Corpus of British English (henceforth FLOB) cited in Brinton (2010) are shown below.

Interpersonal function (speaker's attitude): *I mean*, that's nothing short of treachery (FLOB).

Textual function (as a repair): “I’ll see you in the morning.” She laughed. “*I mean*, afternoon” (FLOB).

Studies on Thai and English Discourse Markers Used by Thai EFL Learners

Some studies have focused on Thai discourse markers used by native Thai speakers (Kittopakrankit, 2018; Simma, 2014; Wutthichamnong, 2016).

Simma (2014) studied the Thai discourse marker /σ∝⊥N/ (ซึ่ง) in Thai discourse and found that it was mostly used as a topic marker to connect utterances. Wutthichamnong (2016) analyzed use of the English loan word *okay* in everyday Thai conversation. In total, 11 pragmatic functions were found to be used by native Thai speakers. In Kittopakrankit’s (2018) research on Thai discourse markers, it was found that Thai discourse markers mainly derived from modal particles used to express politeness and intimacy.

Other studies have focused on the use of English discourse markers in writing by Thai EFL learners (Chotiros, 1999; Jangarun & Luksaneeyanawin, 2016; Prommas & Sinwongsuwat, 2011; Sitthirak, 2013).

Sitthirak (2013) investigated contrastive discourse markers (CDMs) used by Thai university students. The result showed that Thai students can choose different CDMs depending on context and situation.

Chotiros (1999) compared the use of English contrastive discourse markers (ECDMs) and Thai contrastive discourse markers (TCDMs) by Thai EFL learners. The research noted a lack of one-to-one correspondence between the CDMs of the two languages, resulting in different usage of CDMs in both languages.

Both Prommas and Sinwongsuwat (2011) and Jangarun and Luksaneeyanawin (2016) compared differences in the use of discourse connectors in argumentative essays between Thai university students and native English speakers. Prommas and Sinwongsuwat (2011) found that Thai EFL learners used discourse connectors more frequently than native English speakers, whereas native English speakers used more types of discourse connectors. Jangarun and Luksaneeyanawin (2016) found that the most-used discourse connector was *and*, the use of which also represented pragmatic functions other than solely addition.

In contrast to studying written discourse markers, both Nookam (2010) and Arya (2020) investigated the use of spoken discourse markers

by Thai university students. It is worth pointing out that both researchers found that the most-used spoken discourse markers were *and*, *okay*, *but* and *so*. However, interactional functions were rarely identified.

As for the present study, this research focuses on six previously identified spoken discourse markers which are used with high frequency by native English speakers (Ajimer, 1997, 2011; Trillo, 2002), but which have not been studied comprehensively from a Thai EFL learner perspective. The six markers are *like*, *so*, *well*, *you know*, *I think* and *I mean*.

This research, then, focuses on two research questions:

1. Is there a significant difference in the use of the six studied English spoken discourse markers between Thai EFL learners and native English speakers?

2. What are the different pragmatic functions of the six studied English spoken discourse markers used by Thai EFL learners compared to native English speakers?

Methodology

Population and samples

This research focused on L1 Thai EFL learners who were raised and educated to the undergraduate level in Thailand. Based on CEFR assessment standards (Council of Europe, 2001), they were expected to have English proficiency between B1-C1. In this study, 30 Thai B1-level EFL learners and 30 Thai C1-level EFL learners from a public university in Thailand were sampled as participants.[1] Each student's level of English proficiency was determined from his or her CU-TEP score, wherein the cut-off points of the B1 level are 35-69 points, while the cut-off points of the C1 level are 99-120 points (Wudthayagorn, 2018).

Research procedure with research instruments

Six instruments were used in this research. They were: face-to-face prepared questions in selected topics, a recording machine, computerized language analysis (CLAN), Santa Barbara Corpus of Spoken

[1] Performance of B1 and C1 level students will be compared, but that comparison is not included in this paper.

American English (SBCSAE), AntConc, and log-likelihood. Each instrument was introduced as part of the following research procedures.

In order to accurately assess participants' conversational ability across a range of general-conversation topics, the top five content types in contemporary commercial English textbooks were selected. They are:

1. Education
2. People
3. Problems and solutions
4. Relationship
5. Travel

Five questions pertaining to each topic were chosen from the textbooks (For a sample of some specific questions in each topic, please see Appendix 1.) Note, however, that conversations were not limited solely to the prepared questions, as answers by participants sometimes extended to other points.

Each participant had an English conversation with the researcher which lasted for approximately 20 minutes. Conversations were recorded using a smartphone device.

After each conversation, the researcher used CLAN to transcribe spoken data into written data in CHAT format for analysis.

A learner corpus, namely the Thai EFL Learners Spoken English Corpus (TELSEC), was built based on the transcription data, as were two sub-corpora: B1-Level Thai EFL Learner Spoken English Corpus (BTELSEC) and C1-Level Thai EFL Learner Spoken English Corpus (CTELSEC).

AntConc 3.5.8 (Macintosh OS X) 2020 was used to present various descriptive data such as raw frequency of discourse marker use by Thai EFL learners, and to provide concordances for function analysis.

In the process of analysis, tagging discourse markers was a complicated and challenging task, for no systematic and comprehensive method to tag discourse markers has been agreed upon so far. For this reason, the researcher and an interrater followed two principles to manually tag the discourse markers investigated in this research. The first principle was to exclude semantic meanings in context using the *Oxford Dictionary of English (3rd Edition)* (2010) as a reference. The second principle was to identify spoken discourse markers through their unique features based on the findings of previous research (Brinton, 2008), with the features including such conceptualization as phonologically

unstressed, syntactically independent, little or no propositional meaning, and so forth.

Santa Barbara Corpus of Spoken American English (SBCSAE) was used as the native English speaker corpus in this research. SBCSAE is part of International Corpus of English (ICE). It comprises naturally recorded spoken English interactions at different locations throughout the whole of the United States of America. It is, thus, a comparatively large corpus and comprises an enormous amount of data pertaining to everyday spoken English among American native-English speakers. Hence, SBCSAE is suitable as a native English speaker corpus for this research.

Regarding significant difference testing, log-likelihood (LL) calculation was used. As was stated in the literature review, this research largely adopted Brinton's framework of functions of discourse markers (2008) to guide the analysis of functions of spoken discourse markers.

Results and Discussions

Basic information about TELSEC and SBCSAE

The learner corpus, TELSEC, has 75,155 tokens in total. For the six spoken discourse markers investigated in this research, 1,300 spoken discourse markers were identified. SBCSAE has 249,000 tokens in total; 7,096 spoken discourse markers were identified and investigated in this research. Table 1 summarizes the use of spoken discourse markers identified in both TELSEC and SBCSAE.

Table 1

Total Spoken Discourse Markers in TELSEC and SBCSAE and the Use of Each

The keyword	Corpus	Raw frequency as a spoken discourse marker	Proportions (%)
like	TELSEC	429	33.0
	SBCSAE	1,469	20.7
so	TELSEC	188	14.5
	SBCSAE	1,743	24.6
well	TELSEC	24	1.8
	SBCSAE	1,394	19.6

you know	TELSEC	90	6.9
	SBCSAE	1,444	20.3
I think	TELSEC	445	34.2
	SBCSAE	498	7.0
I mean	TELSEC	124	9.5
	SBCSAE	548	7.7
total	TELSEC	1,300	100
	SBCSAE	7,096	100

Significant Differences in the Use of Spoken Discourse Markers Between Thai EFL Learners and Native English Speakers

To examine significant differences, a comparison was made between TELSEC and SBCSAE regarding the use of each spoken discourse marker and the use of all six discourse markers.

Table 2 shows the LL test result from the comparison of all six discourse markers between the two corpora retrieved from LL test by Hardie ($p<0.0001$, $df=6$).

Table 2

LL Test Result from the Comparison of All the Six Spoken Discourse Markers Between TELSEC and SBCSAE

	Value	df	p-value
LL	1414.14	6	0.0000

Table 3 summarizes the LL test results from the comparison of the use of each spoken discourse marker between TELSEC and SBCSAE retrieved from LL Wizard by Rayson ($p<0.0001$, $df=1$) with the LL critical value as 15.13.

In the table, O1 and O2 show the raw frequency in each corpus. %1 and %2 values show relative frequencies in both corpora. The symbol + indicates overuse in O1 relative to O2, while the symbol – indicates underuse in O1 relative to O2.

Table 3

LL Test Results from the Comparison of the Use of Each Spoken Discourse Marker Between TELSEC and SBCSAE

	O1	%1	O2	%2		LL
like	429	0.57	1469	0.59	-	0.36
so	188	0.25	1743	0.70	-	236.18 *
well	24	0.03	1394	0.56	-	562.17 *
you know	90	0.12	1444	0.58	-	339.80 *
I think	445	0.59	498	0.20	+	259.30 *
I mean	124	0.16	548	0.22	-	8.90

As the results show, there was a significant difference in the use of all six spoken discourse markers between Thai EFL learners and native English speakers (LL=1414.14, p=0.0000). As illustrated by the descriptive data and the LL test result, it can be concluded that Thai EFL learners underused English spoken discourse markers compared to native English speakers. This result aligns with some previous research, such as Aijmer (2016), Polat (2011), Buysse (2012), Diskin (2017), et al., reinforcing the notion that non-native English speakers and EFL learners underuse spoken discourse markers in oral communication.

According to the LL test results for each spoken discourse marker, there was no significant difference found for two markers, namely *like* and *I mean*, while a significant difference was found between TELSEC and SBCSAE for the other four spoken discourse markers (*so*, *well*, *you know*, *I think*). Three of them were underused (*so*, *well*, *you know*), while the spoken discourse marker *I think* was overused by Thai EFL learners.

The discourse marker *well* had the largest significant difference. This result echoes that of Nookam (2010), in which no use of *well* was found. The result is also in line with the findings of Fung and Carter (2007), who found that Hong Kong EFL learners underused the discourse marker *well*. Meanwhile, Polat (2011) also produced the same result, confirming that non-native English speakers did not use *well* as a discourse marker in oral communication. Based on previous research, *well* is seen as a focal discourse marker that provides an interpersonal function mainly in oral communication (Aijmer, 2011, 2016; Brinton, 2008). It is thus suggested that Thai EFL learners have a deficiency in using *well* as a spoken discourse marker in English conversation.

A relatively large significant difference was also found for the discourse marker *you know*. Based on previous research, the discourse marker *you know* is a spoken discourse marker that serves various interpersonal functions in oral communication (Aijmer, 2011, 2016). The result may indicate that Thai EFL learners lack the necessary awareness of its use needed for building the interpersonal relationship in conversational interaction.

The LL test result and the descriptive data of the spoken discourse marker *so* show that Thai EFL learners underused the marker in comparison to native English speakers. This result aligns with the study of Buysse (2012), wherein EFL learners seldom used the discourse marker *so* in oral communication.

In contrast to the results above, the spoken discourse marker *I think* was overused by Thai EFL learners compared to native English speakers; in fact, *I think* was the only spoken discourse marker under investigation in this research that was overused by Thai EFL learners. This phenomenon indicates that Thai EFL learners tend to over-rely on use of the spoken discourse marker *I think* in English conversation.

In conclusion, the results illustrate that Thai EFL learners generally underused spoken discourse markers. Of the six spoken discourse markers, *well*, *so* and *you know* were underused by Thai EFL learners compared to native English speakers, with the largest difference being in use of the spoken discourse marker *well*. By contrast, the spoken discourse marker *I think* was overused by Thai EFL learners.

Different Pragmatic Functions of Spoken Discourse Markers Used by Thai EFL Learners Compared to Native English Speakers

All pragmatic functions of each spoken discourse marker in TELSEC and SBCSAE were identified based on Brinton's (2008) framework of the pragmatic functions of discourse markers and the findings related to pragmatic functions in previous research. Based on the identified functions in both corpora, differences in pragmatic functions of each spoken discourse marker were derived from three facets: the normalized frequency of each function of each spoken discourse marker, the LL test result of each function of each spoken discourse marker, and the proportion of the participants who used each function of each spoken discourse marker. Focus was placed on these three facets with the

understanding that they would provide comprehensive information for purpose of comparing spoken discourse markers of the two corpora.

In this section, differences in the use of each spoken discourse marker, as well as all six spoken discourse markers, will be presented in succession. Moreover, only those functions of each spoken discourse marker evincing major differences between the two corpora are presented in the tables.

In all the tables below, NF stands for the normalized frequency of each function (per 10,000 tokens); NP stands for the number of participants who used the corresponding function; P stands for proportion; TF stands for textual function; IF stands for interpersonal function.

Table 4 gives details about three functions of the spoken discourse marker *like* in which major differences were found between the two corpora.

Table 4

Three Functions of The Spoken Discourse Marker 'like' in Which Major Differences Were Found Between TELSEC and SBCSAE

like	SBCSAE	TELSEC	LL	SBCSAE		TELSEC	
Function (all are TFs)	NF	NF		NP	P (%)	NP	P (%)
Searching for the right words	15	2	- 116.77	205	68.8	4	6.7
To exemplify	16	15	- 0.03	182	61.1	16	26.7
Marking a focus on new information	13	25	+ 49.12	132	44.3	25	41.7

From the table above, a substantial difference can be seen with the function, *searching for the right words*; a limited number of Thai EFL learners used it at low frequency, while a large number of native English speakers used it at a much higher frequency. This result suggests that Thai EFL learners lack an awareness of using this function in conversation. Meanwhile, the proportion of Thai EFL learners who used the function, *to exemplify*, was less than half that of native English speakers. Though there was no significant difference in the normalized frequencies of using this function between the two corpora.

Conversely, an interesting finding is that the function of *like* used by Thai EFL learners with the highest frequency was *marking a focus on new information*, which was different from the most common functional use by native English speakers, namely, *to exemplify*. Since the proportion of the participants who used the function*, marking a focus on new information*, was almost the same in both corpora, it is suggested that Thai EFL learners tend to stick to using this function in conversation.

Similar to the findings of Diskin (2017), non-native English speakers used the spoken discourse marker *like* with function, to exemplify, at a relatively high frequency. Nonetheless, it is likewise interesting to note that the function, *more information or explanation*, was used at the highest frequency in Diskin (2017), while it was used at relatively low frequency by Thai EFL learners. This dissimilarity may indicate different ways of using spoken discourse markers by different groups of non-native English speakers.

Table 5 provides details about two functions of the spoken discourse marker *so* in which major differences were found between the two corpora.

Table 5

Two Functions of the Spoken Discourse Marker 'so' in Which Major Differences Were Found Between TELSEC and SBCSAE

so	SBCSAE	TELSEC	LL	SBCSAE		TELSEC	
Function	NF	NF		NP	P (%)	NP	P (%)
Marking a question (IF)	15	0	-169.05	217	72.8	3	5.0
Marking a result/consequence (TF)	12	9	-5.68	170	57.0	11	18.3

As Table 5 shows, the biggest difference regarding use of *so* between the two corpora was the interpersonal function, *marking a question*. Only three Thai EFL learners were found to use it, while native English speakers used it with comparatively high frequency. On the other hand, Thai EFL learners focused on two textual functions, *marking a result or consequence* and *marking a start of a new narration*. As shown in the table, the proportion of Thai participants who used the function,

marking a result or consequence was only one-third that of native English speakers, although normalized frequencies were similar and without significant difference. This indicates that a limited range of Thai EFL learners used this function, yet, those who did use the function placed substantial focus on it. Such findings square with those of Nookam (2010) and Arya (2020), indicating that Thai EFL learners emphasize fewer interpersonal functions in interactions, while instead, giving focus to textual functions.

Furthermore, based on the data, 10 out of 12 functions are associated with a phenomenon whereby the proportion of native English speakers who used the functions was two to seven times higher than the proportion of Thai EFL learners who did so, with five of the functions being textual and the other five interpersonal. This indicates that only a limited number of Thai EFL learners used *so* as a discourse marker in communication.

A notable finding pertains to the spoken discourse marker *well* and differences between SBCSAE and TELSEC. Only 24 cases of its use were found in TELSEC. Table 6 summarizes all interpersonal functions and textual functions of *well* that were found in SBCSAE but not found in TELSEC.

Table 6

All the Interpersonal Functions and Textual Functions of the Spoken Discourse Marker 'well' That Were Found Only in SBCSAE

well	SBCSAE	TELSEC		SBCSAE		TELSEC	
Function	NP	NP	LL	NP	P (%)	NP	P (%)
All the IFs found in SBCSAE, but missing in TELSEC	19	0	-243.19	177	59.4	0	0
All the TFs found in SBCSAE, but missing in TELSEC	5	0	-41.96	100	33.6	0	0

As seen in Table 6, Thai EFL learners used neither the interpersonal functions nor the textual functions, including *face-threat mitigator*, *marking a request*, *marking a question* and other functions. In fact, the data reveal that a very limited number of Thai EFL learners – only three out of 60 – used the spoken discourse marker *well*, resulting in

a large discrepancy between participants who used *well* in English conversation in the two corpora.

Although this finding regarding the use of *well* is in line with Nookam (2010), as mentioned in the previous section, it differs from some other research (Aijmer, 2011; Müller, 2005) in which EFL learners were found to use some functions of *well* more than native English speakers, including the functions of *searching for the right words*, *as a repair*, and other functions. The results may indicate substantial differences regarding the use of *well* by EFL learners from different backgrounds.

Also deserving mention is a point on the spoken discourse marker *well* and its significant role in determining the relationship between speaker and hearer. This is due to the word's interpersonal functions which, should attract more study than its textual functions as presented in previous research (Aijmer, 2011; Sakita, 2013). Therefore, the low performance of *well* in TELSEC indicates that Thai EFL learners were notably deficient in detecting relevant connections between speaker and hearer in conversational interactions.

Similar to the discourse marker *well*, infrequent use of the discourse marker *you know* observed among Thai EFL learners coincided with lack of use of various interpersonal functions and textual functions such as *as a repair*, *as a mitigator*, and the like, while native English speakers' use of *you know* ranged across a wider set of functions, as illustrated in Table 7.

Table 7

Major Differences in the Use of Five Functions of the Spoken Discourse Marker 'you know' Between TELSEC and SBCSAE

You know	SBCSAE	TELSEC		SBCSAE		TELSEC	
Function	NF	NF	LL	NP	P (%)	NP	P (%)
Marking an involvement (IF)	12	6	-21.95	141	47.3	4	6.7
Searching for the right words (TF)	7	0	-79.30	129	43.3	1	1.7
All the other TFs found in SBCSAE, but missing in TELSEC	5	0	-55.16	64	21.5	0	0

All the other IFs found in SBCSAE, but missing in TELSEC	4	0	-50.56	103	34.6	0	0
An implicature of common knowledge (IF)	4	3	-1.14	77	25.8	4	6.7

In line with results depicted in Table 7, Thai EFL learners focused on using only two functions: *marking an involvement* and *an implicature of the shared or common knowledge*, with use frequency of the function, *marking an involvement*, remaining significantly lower than native English speakers. House (2009) found that non-native English speakers often did not consider the interpersonal purpose of using *you know*. Moreover, more than half of native English speakers in SBCSAE used *you know,* while only five Thai participants used it. It is therefore suggested that Thai EFL learners may lack understanding regarding the use both types of functions of the spoken discourse marker *you know,* based on the low performance of it.

Table 8 summarizes major differences in the use of two functions of the spoken discourse marker *I think* between the two corpora.

Table 8

Major Differences in the Use of Two Functions of the Spoken Discourse Marker 'I think' Between TELSEC and SBCSAE

I think	SBCSAE	TELSEC		SBCSAE		TELSEC	
Function	NF	NF	LL	NP	P (%)	NP	P (%)
An implicature of opinion/feeling/stance (IF)	14	54	+341.33	245	82.2	60	100
As a mitigator (IF)	3	0	-27.34	61	20.5	2	3.3

As reflected in Table 8, an over-reliance on the function, *an implicature of opinion, feeling or stance*, of the spoken discourse marker *I think* was found among Thai EFL learners in TELSEC. In contrast, other functions including, *as a mitigator*, were rarely performed by Thai EFL learners. This result indicates that Thai EFL learners stick to using only one function of the spoken discourse marker *I think*. This can be construed as an inappropriate use of *I think,* owing to an over-reliance on one specific function, as argued in previous research (Aijmer, 2011).

Table 9 summarizes the major differences in the use of four functions of the spoken discourse marker *I mean* between the two corpora.

Table 9

Major Differences in the Use of Four Functions of the Spoken Discourse Marker 'I mean' found in TELSEC and SBCSAE

I mean	SBCSAE	TELSEC	LL	SBCSAE		TELSEC	
Function	NF	NF		NP	P (%)	NP	P (%)
An implicature of the cause subjectively (IF)	5	2	-14.15	115	38.6	4	6.7
As a mitigator (IF)	4	0	-40.40	88	29.5	0	0
Searching for the right word (TF)	2	0	-18.48	45	15.1	1	1.7
As a repair (TF)	1	4	+27.20	17	5.7	4	6.7

As can be seen from Table 9, native English speakers used the function, *an implicature of the cause subjectively*, at the highest frequency, while Thai EFL learners focused primarily on the function, *as a repair*. No performance of the function*, as a mitigator*, was found in TELSEC, whereas a large number of native English speakers used it very frequently. The function, *searching for the right word*, was underused by Thai EFL learners. Another function of the discourse marker *I mean*, namely, *as a repair*, the function of which has been argued as deriving from its semantic meaning (Brinton, 2010), was used much more often by Thai EFL learners than by native English speakers. Furthermore, a larger number of native English-speaking participants used *I mean* than did Thai EFL learners, illustrating that a narrower range of Thai EFL learners used the spoken discourse marker *I mean* in English conversation.

Overall, as the comparisons above indicate, a central finding is that Thai EFL learners underused both types of functions compared to native English speakers, while a more widespread deficiency in using interpersonal functions, as opposed to textual functions, was also evident; see Table 10 below. It is therefore conceivable that discrepancy

in performance of interpersonal functions of spoken discourse markers is larger than any discrepancy relating to performance of textual functions by Thai EFL learners.

Table 10

Differences in the Use of Both Types of Functions of All the Six Spoken Discourse Markers Between TELSEC and SBCSAE

Six spoken discourse markers	SBCSAE	TELSEC	LL	SBCSAE		TELSEC			
Function	NF	NF		NP	P (%)	NP		P (%)	
All the textual functions	136	92	-97.65	257	86.2	32		53.3	
All the interpersonal functions	137	80	-164.22	265	88.9	60	14	100	23.3

As shown in Table 10, there was a large difference in the number of participants who used both types of functions between the two corpora.

Here, it is worth mentioning a point concerning the number of participants who used interpersonal functions in TELSEC. Since all Thai participants in this research used the function, namely, *an implicature of opinion, feeling or stance*, of the spoken discourse marker *I think*, under the circumstances it was considered inappropriate to show merely the number 60 (100%). To illustrate the condition of using interpersonal functions more objectively and comprehensively, another number 14 (23.3%) is presented in the table, illustrating that only 14 out 60 Thai participants (23.3%) used the interpersonal functions in the situation where the function 'an implicature of opinion, feeling or stance' of *I think* was excluded.

These findings are in line with several previous research works (Arya, 2020; Asik & Cephe, 2013; Fung & Carter, 2007; Nookam, 2010; Polat, 2011; Trillo, 2002) in which EFL learners were found to use fewer functions compared to native English speakers. In particular, Arya (2020) and Nookam (2010) identified a general deficiency in using spoken

discourse markers in English conversation. Moreover, comparisons of functions of spoken discourse markers in this research reinforce the notion, as pointed out by Arya (2020) and Nookam (2010), that Thai EFL learners lack the use of interpersonal functions in terms of person-to-person interactions. It is likewise worth pointing out that, in both Arya (2020) and Nookam (2010), the discourse marker *so* was found to be one of the most-used spoken discourse markers by Thai EFL learners. Despite this, it can be seen that the spoken discourse marker *so* was, in fact, underused by Thai EFL learners compared to native English speakers.

There may be reason to consider the influence of L1 Thai with respect to the lack of use of English spoken discourse markers by Thai EFL learners. With previous research in mind, it was found that native Thai speakers use Thai spoken discourse markers for politeness and intimacy (Kittopakrankit, 2018) rather than for interpersonal purposes, with cultural and social impact being significant factors here. Moreover, as previously suggested (Chotiros, 1999), discourse markers in both languages may be without one-to-one correspondence, making it difficult for Thai EFL learners to acquire English spoken discourse markers.

In light of the examinations of discourse markers addressed in previous research (Aijmer, 2011; Fung & Carter, 2007; House, 2013), it can be seen that multi-functional spoken discourse markers affect oral communication across various aspects from a pragmatic perspective, such as, for instance the interpersonal relationship between interlocutors, or situational understanding between speaker and hearer, among other aspects. Hence, this entire category is regarded as a focal component of expressing pragmatic meaning intended by native English speakers, and is an important acquisition for EFL learners in the process of becoming competent English speakers.

Tracing back, then, to findings of this research, an interpretation of data reveals that, to some extent, Thai EFL learners lack pragmatic competence in oral communication in terms of discrepancies in the performance of spoken discourse markers compared to native English speakers. Interpretation further reveals, when considered together with previous research (Fung & Carter, 2007), that Thai EFL learners may be exposed to unnatural linguistic input in traditional learning environments, specifically a focus on English grammar and propositional meanings of vocabulary, thus resulting in low frequency use of spoken discourse markers in general. It should be noted that a normative claim that EFL

learners should or must use spoken discourse markers in the same way or at a similar frequency as native English speakers does not inhere in previous research. However, taking into account structure and function as used in utterance, some spoken discourse markers are more difficult to acquire by non-native English speakers (Aijmer, 2011; Diskin, 2017) and require a high level of ability in terms of linguistics, socio-pragmatic awareness, etc. among EFL learners. Hence, the general instance of low performance of English spoken discourse markers by Thai EFL learners further reflects learners' unstable and incomplete acquisition of English as a foreign language (Diskin, 2017).

Conclusion

On the whole, Thai EFL learners underused a majority of pragmatic functions of English spoken discourse markers investigated in this research (59% of pragmatic functions were underused). In one regard, such underuse of spoken discourse markers may be considered a general phenomenon among non-native English speakers inasmuch as they produce the necessary conditions for marker use less frequently in English conversation than do native English speakers. Otherwise, an assumption is made that performing English spoken discourse markers for purpose of achieving pragmatic competence in communication is not an easy task for EFL learners.

In addition, it is posited that Thai EFL learners exhibit deficiency in performing interpersonal functions of spoken discourse markers in oral communication, particularly regarding the two interpersonal-centered spoken discourse markers *well* and *you know*.

Taking a cue from the findings above, including the stated infrequent use of various functions of each spoken discourse marker by participants, as well as the low distribution of same among them, the present researchers are led to the fact that Thai EFL learners exhibit deficiency in pragmatic competence in English oral communication. Likewise emerging from the findings is that such deficiency owes in part to insufficient acquisition of English spoken discourse markers, as well as the influence of L1 Thai language on marker use.

Accordingly, this research suggests that Thai EFL learners should firstly acquire a general concept of English spoken discourse markers, including relevant definitions, features and so forth. This may help Thai

EFL learners to better understand the importance of English spoken discourse markers in English conversation and to become more aware of when and how they are used. Furthermore, EFL learners should acquire the aforementioned general concept in a systematic way that takes in the main pragmatic functions of English spoken discourse markers. In this light, and exemplifying the importance of effective concept acquisition, the two spoken discourse markers *well* and *you know* may require increased focus, given the considerable deficiency in using them identified in the research.

Indeed, additional English spoken discourse markers used by Thai EFL learners should be studied comprehensively in order to examine patterns and adequately determine how Thai EFL learners use the markers in English oral communication. Moreover, further comparisons of Thai spoken discourse marker and English spoken discourse marker use should be studied with aim of investigating how L1 Thai language influences the use of English spoken discourse markers by Thai EFL learners. Lastly, it must be noted that the findings of this research are based solely on comparisons to American English as recorded in SBCSAE. Different findings may, and most likely will, occur should comparisons be made with other English dialects.

Acknowledgements

I would like to express my deepest appreciation to my advisor, Associate Professor Wirote Aroonmanakun, Ph.D., who has been patient in guiding me and supervising my work all along the way. Moreover, I am grateful to all the participants who were willingly to be involved in this research.

About the Authors

Zhaoyi Pan: Ph.D. candidate in the English as an International Language Program, Graduate School, Chulalongkorn University, Thailand. His research interests are corpus linguistics and pragmatics.

Wirote Aroonmanakun: Associate professor in the Department of Linguistics; Corpus Linguistics for Digital Humanities Research Unit, Faculty of Arts, Chulalongkorn University. His research interests are corpus linguistics and computational linguistics.

References

Adams, C. (2002). Practitioner review: The assessment of language pragmatics. *Journal of Child Psychology and Psychiatry*, *43*, 973–987. https://doi.org/10.1111/1469-7610.00226

Aijmer, K. (2004). Pragmatic markers in spoken interlanguage. *Nordic Journal of English Studies, 3*(1), 173-190.

Aijmer, K. (2011). Well I'm not sure I think...The use of well by non-native speakers. *International Journal of Corpus Linguistics, 16*(2), 231-254. https://doi.org/10.1075/ijcl.16.2.04aij

Aijmer, K. (2016). Pragmatic markers as constructions. The case of anyway. In G. Kaltenbock, E. Keizer & A. Lohmann (Eds.), *Outside the clause* (pp. 29-57). John Benjamins Publishing Company.

Anthony, L. (2019). *AntConc (Version 3.5.8) [Computer Software]*. http://www.laurenceanthony.net/software

Arya, T. (2020). Exploring discourse marker use in Thai university students' conversations. *LEARN Journal: Language Education and Acquisition Research Network Journal*, *13*(1), 247-267.

Asik, A. & Cephe, P. T. (2013). Discourse markers and spoken English: Nonnative use in the Turkish EFL setting. *English Language Teaching, 6*(12), 144-155. http://dx.doi.org/10.5539/elt.v6n12p144

Bazzanella, C. (2006). Discourse markers in Italian: Towards a "Compositional" Meaning. *Kerstin Fischer,* 449-464.

Bignell, S. & Cain, K. (2007). Pragmatic aspects of communication and language comprehension in groups of children differentiated by teacher ratings of inattention and hyperactivity. *British Journal of Developmental Psychology, 25*, 499-512.

Bolden, G. B. (2015). Discourse markers. In K. Tracy (Ed.), *The international encyclopedia of language and social interaction* (pp. 1-8). John Wiley & Sons, Inc.

Brinton, L. J. (2008). *The comment clause in English: Syntactic origins and pragmatic development*. Cambridge University Press.

Brinton, L. J. (2010). The development of *I mean*: Implications for the study of historical pragmatics. In S. M. Fitzmaurice & I. Taavitsainen (Eds.). *Methods in historical pragmatics* (pp. 37–80). Mouton de Gruyter.

Buysse, L. (2012). *So* as a multifunctional discourse marker in native and learner speech. *Journal of Pragmatics, 44*, 1764-1782. https://doi.org/10.1016/j.pragma.2012.08.012

Chotiros, P. (1999). *A comparative analysis of Thai and English contrastive discourse markers: With a discussion of the pedagogical implications* [Unpublished doctoral dissertation]. Boston University.

CLARINO Bergen Centre. (1991). *The Freiburg-LOB Corpus of British English (FLOB)*. Retrieved from http://clarino.uib.no/korpuskel/landing-page?identifier=flob&view=short

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Structured overview of all CEFR scales*. Available at https://rm.coe.int/168045b15e.

Diskin, C. (2017). The use of the discourse-pragmatic marker "like" by native and non-native speakers of English in Ireland. *Journal of Pragmatics, 120*, 144-157. https://doi.org/10.1016/j.pragma.2017.08.004

Fraser, B. (1988). Types of English discourse markers. *Acta Linguistica Hungarica, 38*(1), 19-33.

Fraser, B. (2011). The sequencing of contrastive discourse markers in English. *Baltic Journal of English language, Literature and Culture, 1*, 29-35.

Fung, L. & Carter, R. (2007). Discourse markers and spoken English: Native and learner use in pedagogical settings. *Applied Linguistics, 28*(3), 410-439. https://doi.org/10.1093/applin/amm030

Hardie, A. *Welcome to the UCREL significance test system*. UCREL Significance Test System. Retrieved March 1, 2022, from http://corpora.lancs.ac.uk/sigtest/

House, J. (2009). Subjectivity in English as lingua franca discourse: The case of *you know*. *Intercultural Pragmatics, 6*(2), 171–193. https://doi.org/10.1515/IPRG.2009.010

House, J. (2013). Developing pragmatic competence in English as a lingua franca: Using discourse markers to express (inter)subjectivity and connectivity. *Journal of Pragmatics*, *59*, 57-67. https://doi.org/10.1016/j.pragma.2013.03.001

Jangarun, K., & Luksaneeyanawin, S. (2016). Discourse connector usage in argumentative essays by American and Thai university students.

Journal of Pan-Pacific Association of Applied Linguistics, 20(1), 95-112.

Kittopakrankit, K. (2018). A comparative study of discourse markers between Chinese and Thai languages. *Chinese Studies Journal, 11*(2), 66-101.

Müller, S. (2005). *Discourse markers in native and non-native English discourse.* John Benjamins Publishing Co.

Nookam, W. (2010). *Thai EFL learners' use of discourse markers in English conversation: A study of business English students at Didyasarin International College* [Unpublished master's thesis]. Prince of Songkla University.

Polat, B. (2011). Investigating acquisition of discourse markers through a developmental learner corpus. *Journal of Pragmatics, 43*(15), 3745-3756. https://doi.org/10.1016/j.pragma.2011.09.009

Prommas, P. & Sinwongsuwat, K. (2011). A comparative study of discourse connectors used in argumentative compositions produced by Thai EFL learners and English-native speakers. *The 3rd International Conference on Humanities and Social Sciences,* 1-17.

Rayson, P., Berridge, D., & Francis, B. (2004, March 10-12). Extending the Cochran rule for the comparison of word frequencies between corpora. *Le poids des mots: Proceedings of the 7th International Conference on Statistical analysis of textual data (JADT 2004)*, Louvain-la-Neuve, Belgium.

Sakita, T. I. (2013). Discourse markers as stance markers. *Pragmatics Cognition, 21*(1), 81-116. https://doi.org/10.1075/pc.21.1.04sak

Schiffrin, D. (1987). *Discourse markers*. Cambridge University Press.

Simma, P. (2014). Function of /σ∝⊥N/ (ซึ่ง) in discourse. *Ramkhamhaeng University Journal Humanities Edition, 33*(1), 1-22.

Sitthirak, C. (2013). A comparison between Thai university students and English speakers using contrastive discourse markers. *Foreign Language Teaching and Learning* (*FLLT) Conference, 2*(1), 875-886.

Stevenson, A. (2010). *Oxford dictionary of English (3rd Edition).* Oxford University Press.

Traugott, E. C. & Dasher, R. B. (2002). *Regularity in semantic change.* Cambridge University Press.

Trillo, J. R. (2002). The pragmatic fossilization of discourse markers in non-native speakers of English. *Journal of Pragmatics, 34*, 769-784. https://doi.org/10.1016/S0378-2166(02)00022-X

Wudthayagorn, J. (2018). Mapping the CU-TEP to the Common European Framework of Reference (CEFR). *LEARN Journal: Language Education and Acquisition Research Network Journal*, *11*(2), 163-180.

Wutthichamnong, W. (2016). Pragmatic functions of "Okay" in Thai Conversation. *Journal of Community Development Research (Humanities and Social Sciences), 9*(2), 115-127.

Appendix 1

Sample questions in the conversations with Thai EFL learners

Question relevant to the topic of "education":
Tell me something about your school life before university.

Question relevant to the topic of "people":
How do you see yourself?

Question relevant to the topic of "problems and solutions":
Was there anything bothered you when you were growing up?

Question relevant to the topic of "relationship at home":
What was your childhood like?

Question relevant to the topic of "travel":
Tell me about a place that you have visited to.