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

Mr. Zhaoyi Pan

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

A Dissertation Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in English as an International Language

Inter-Department of English as an International Language

GRADUATE SCHOOL

Chulalongkorn University

Academic Year 2021

การใช้ดัชนีปริจเฉทภาษาพูดของผู้เรียนภาษาอังกฤษชาวไทย: การศึกษาแบบอิงคลังข้อมูลผู้เรียน

นายเจ้ายี่ แพน

Thesis Title	A Learner Corpus-Based Study on the Use of Spoken Discourse Markers by Thai EFL Learners
By	Mr. Zhaoyi Pan
Field of Study	English as an International Language
Thesis Advisor	Associate Professor WIROTE AROONMANAKUN, Ph.D.

Accepted by the GRADUATE SCHOOL, Chulalongkorn University in Partial Fulfillment of the Requirement for the Doctor of Philosophy

.... Dean of the GRADUATE SCHOOL

(Associate Professor YOOTTHANA CHUPPUNNARAT, Ph.D.)

DISSERTATION COMMITTEE

.... Chairman

(Associate Professor Supakorn Phoocharoensil, Ph.D.)

.... Thesis Advisor

(Associate Professor WIROTE AROONMANAKUN, Ph.D.)

.... Examiner

(Associate Professor NATTAMA PONGPAIROJ, Ph.D.)

.... Examiner

(Associate Professor RAKSANGOB WIJITSOPON, Ph.D.)

.... Examiner

(Assistant Professor THEERAPORN RATITAMKUL, Ph.D.)

iii

เจ้ายี่ แพน : การใช้ดัชนีปริจเฉทภาษาพูดของผู้เรียนภาษาอังกฤษชาวไทย: การศึกษาแบบอิง
คลังข้อมูลผู้เรียน. (A Learner Corpus-Based Study on the Use of Spoken Discourse Markers by
Thai EFL Learners) อ.ที่ปรึกษาหลัก : รศ.วิโรจน์ อรุณมานะกุลดร.

งานวิจัยนี้ศึกษาและเปรียบเทียบการใช้ดัชนีปริจเฉทภาษาพูดภาษาอังกฤษ ของผู้เรียนชาวไทยกับผู้
พูดภาษาอังกฤษเป็นภาษาแม่ ซึ่งมีทั้งสิ้นหกดัชนี ได้แก่ *like, so, well, you know, I think,* และ *I mean* โดย
เปรียบเทียบในสองด้าน ได้แก่ ความถี่และหน้าที่ทางวัจนปฏิบัติศาสตร์ มีผู้เรียนชาวไทยเข้าร่วมงานวิจัยนี้
ทั้งสิ้น 60 คนแบ่งเป็นผู้เรียนภาษาอังกฤษชาวไทยที่มีความสามารถทางภาษาอังกฤษระดับ B1 และ C1 ตาม
กรอบมาตรฐาน CEFR จำนวนอย่างละ 30 คน ข้อมูลภาษาพูดถูกรวบรวมผ่านการสนทนาในหัวเรื่องเกี่ยวกับ
ชีวิตประจำวันและถอดความเป็นภาษาเขียนเพื่อสร้างคลังข้อมูลผู้เรียนเสำหรับใช้วิเคราะห์ อีกทั้งงานวิจัยนี้ยัง
ได้ศึกษาความแตกต่างอย่างมีนัยสำคัญของการใช้ดัชนีปริจเฉทภาษาพูดภาษาอังกฤษระหว่างผู้เรียน
ภาษาอังกฤษชาวไทยและผู้พูดภาษาอังกฤษเป็นภาษาแม่ นอกจากนั้น งานวิจัยนี้ยังได้ระบุและเปรียบเทียบ
หน้าที่ทางวัจนปฏิบัติศาสตร์ของปริจเฉทภาษาพูดภาษาอังกฤษแต่ละคำที่ใช้โดยผู้เรียนภาษาอังกฤษชาวไทย
กับผู้พูดภาษาอังกฤษเป็นภาษาแม่อีกด้วย

ผลการศึกษาพบความแตกต่างอย่างมีนัยสำคัญในการใช้ดัชนีปริจเฉทภาษาพูดสี่ดัชนี
ได้แก่ *so, well, you know,* และ *I think* ที่ผู้เรียนภาษาอังกฤษชาวไทยใช้น้อยกว่า นอกจากนี้ ผู้เรียน
ภาษาอังกฤษชาวไทยยังใช้ดัชนีปริจเฉทภาษาพูดแต่ละดัชนีแตกต่างกันไปเมื่อเทียบกับผู้พูดภาษาอังกฤษเป็น
ภาษาแม่ โดยภาพรวม หน้าที่ด้านบุคคลสัมพันธ์พบใช้น้อยกว่าหน้าที่ด้านตัวบท ซึ่งแสดงว่าผู้เรียน
ภาษาอังกฤษชาวไทยมีความบกพร่องด้านการใช้ดัชนีปริจเฉทเพื่อหน้าที่บุคคลสัมพันธ์ ผลการศึกษาทำให้
สรุปได้ว่าเมื่อเทียบกับผู้พูดภาษาอังกฤษเป็นภาษาแม่ ผู้เรียนภาษาอังกฤษชาวไทยมีความสามารถทางวัจน
ปฏิบัติศาสตร์ในด้านการใช้ดัชนีปริจเฉททำหน้าที่บุคคลสัมพันธ์ได้ไม่เท่ากับผู้พูดภาษาอังกฤษเป็นภาษาแม่

สาขาวิชา	ภาษาอังกฤษเป็นภาษานานาชาติ	ลายมือชื่อนิสิต ....
ปีการศึกษา	2564	ลายมือชื่อ อ.ที่ปรึกษาหลัก ....

6187758220 : MAJOR ENGLISH AS AN INTERNATIONAL LANGUAGE

KEYWORD: spoken discourse markers, pragmatic functions, Thai EFL learners, learner corpus

Zhaoyi Pan : A Learner Corpus-Based Study on the Use of Spoken Discourse Markers by Thai EFL Learners. Advisor: Assoc. Prof. WIROTE AROONMANAKUN, Ph.D.

This research investigated and compared the use of English spoken discourse markers by Thai EFL learners and by native English speakers. Six English spoken discourse markers, namely *like*, *so*, *well*, *you know*, *I think* and *I mean,* are compared in two aspects: 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 according to the CEFR standard. Spoken data in the genre of English daily conversation was collected and transcribed into written form to build a learner corpus for the analysis. The research examined the significant difference in the use of English spoken discourse markers between Thai EFL learners and native English speakers. Meanwhile, pragmatic functions of each spoken discourse marker used by Thai EFL learners were identified and compared to native English speakers.

The data and its analysis indicated a significant difference in the use of four spoken discourse markers – *so*, *well*, *you know* and *I think* – resulting in underuse by Thai EFL learners. Moreover, Thai EFL learners proved to use each spoken discourse marker differently in comparison to native English speakers. On the whole, interpersonal functions were a factor less frequently 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 use discrepancy regarding spoken discourse markers compared to native English speakers.

| Field of Study: | English as an International Language | Student's Signature .... |
| Academic Year: | 2021 | Advisor's Signature .... |

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

LIST OF TABLES

XV

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

LIST OF FIGURES

Chapter 1

Introduction

This chapter presents the background of this research, the rationale of this research, the research questions, the objectives of this research, the main hypotheses, the scope of this research, the important definitions, and the significance of this research. Each part will be illustrated one after another as follows.

1.1 Background of the study

Based on Fraser's (1988) early introduction of the existence of discourse markers in communication, research on discourse markers has been studied for over 40 years. Fraser (1988) pointed out that one of the earliest researchers who started to study discourse markers was Levinson (1983), even though Levinson had not presented the term "discourse markers" at that moment. It was illustrated by Bolden (2015) that since the late 20th century, discourse markers have been studied by different researchers in various contexts. Through the previous research on discourse markers in both spoken and written forms, what makes discourse markers so important to study became an interesting topic to argue and discuss.

Early researchers studied discourse markers in both written form and spoken form used by native English speakers (Aijmer, 1997, 2004; Fraser, 1988, 1993, 1996, 1999; Levinson, 1983; Schiffrin, 1987). What they found first was that discourse markers were important in connecting discourses or utterances in communication. Fraser (1988, 1993, 1996, 1999) investigated several discourse markers, i.e., *but*, *so*, *you know*, etc., and pointed out that they were as connectors to link utterances, where they were used in the middle of one foregoing utterance and one up-coming utterance. Thus, the mode is "S1+DM(s)+S2", wherein they signal a relationship between the segment they introduce, S2, and the prior segment, S1. In recent research, this

function has also been discovered within different languages (Aull, 2019; Povolna, 2012; Rezanova & Kogut, 2015; Waltereit & Detges, 2007). It is thus believed that the basic function of a discourse marker is to link utterances.

With more studies, the research on discourse markers has been separated into written discourse markers and spoken discourse markers. For the written discourse markers, previous research tended to study the written discourse connectors such as *and*, *so*, etc. (Babanoglu, 2014; Modhish, 2012; Rahimi, 2011) and focused on contrastive connectors such as *but* and *yet* (Povolna, 2012).

Besides connecting utterances, functions of written discourse markers have always interested different researchers, wherein Coherence Theory was adopted to analyze different functions (Fraser, 1988; Schiffrin, 1985). Fraser (1988) explained three types of relations in the functions of written discourse markers: the topic shift relation, the activity relation, and the message relation. Each type also has sub-types which will be presented in the literature review in detail.

At the same time, studies on spoken discourse markers were also conducted. These studies found that spoken discourse markers were different from written discourse markers (Aijmer, 2016; Bolden, 2015). Even though the basic function of a discourse marker was to connect discourses, spoken discourse markers surely had much more pragmatic functions than that (Aijmer, 2016). Bolden (2015) pointed out the fact that discourse markers in spoken language may have their valuable pragmatic functions. It was also aligned with the studies by Aijmer (2016) and Borderia (2018). Aijmer (2016) has kept studying discourse markers in spoken English and explained that discourse markers were unique in communication. For one thing, they appeared in oral communication at a high frequency (Aijmer, 2016; Bolden, 2015). For another, their pragmatic functions played important roles in oral communication (Borderia, 2018; House, 2013). They should be discovered and studied in conversations.

2

Borderia (2018) also argued that discourse markers belonged to oral communication. They were found everywhere in conversations and speech forms. Recent experimental studies on the pragmatic functions of discourse markers (Buysse, 2012; Diskin, 2017; Fung & Carter, 2007; House, 2013; Polat, 2011) further proved that discourse markers had their uniqueness in oral communication, and their valuable pragmatic functions could only be examined in oral communication, especially in conversations (Aijmer, 2016; Blakemore &Gallai, 2014).

As is known, communication is full of pragmatics (Bignell & Cain, 2007). Pragmatic competence is one of the most important competence in communication (Adams, 2002). Due to the uniqueness of the functions of each spoken discourse marker, the speaker and the hearer may continuously understand each other pragmatically by using various spoken discourse markers in conversation. Thus, one of the significant aspects of studying spoken discourse markers was to examine their unique pragmatic functions in communication (Aijmer, 2004, 2011). Under the circumstances, previous research tended to study the pragmatic functions of spoken discourse markers produced by native English speakers. Given that each spoken discourse marker had its unique functions, different researchers tried to examine the pragmatic functions of each spoken discourse marker specifically.

Recent studies focused on comparisons between native speakers and non-native speakers (Aijmer, 2004; Buysse, 2012; Diskin, 2017; Fung & Carter, 2007; House, 2013; Polat, 2011). Most of the research collected non-native English speakers' data to compare with the data of native English speakers in the same genre. Basically, in previous research, two major problems in the use of spoken discourse markers by non-native English speakers were discovered.

The first problem was the different frequencies of using the spoken discourse markers between non-native English speakers and native English speakers. Some

3

studies discovered that the frequency of using the spoken discourse markers of non-native English speakers was lower than native English speakers (Aijmer, 2004; Hellerman & Vergun, 2007), while some studies discovered that non-native English speakers used certain spoken discourse markers in a higher frequency than native English speakers (Fung & Carter, 2007; Polat, 2011). Meanwhile, some studies found that non-native English speakers were not aware of using the spoken discourse markers in conversations so that no spoken discourse marker was produced (Aijmer, 2004; Polat, 2011).

The second problem was the different uses of some functions of the spoken discourse markers between native English speakers and non-native English speakers. It was found that non-native English speakers used fewer functions of spoken discourse markers compared to the functions used by native English speakers (Diskin, 2017; Polat, 2011), while some studies discovered that EFL learners seemed to have problems understanding how to use certain spoken discourse markers (Aijmer, 2004; Trillo, 2002).

To sum up, it can be seen that research on different EFL learners with different backgrounds may lead to different results, either an underuse or an overuse of the spoken discourse markers. Meanwhile, non-native English speakers had problems using different functions of spoken discourse markers, leading to failure or misunderstanding in communication (Bolden, 2006, 2009; Fung & Carter, 2007).

Previous research used EFL learners with different backgrounds, such as Chinese English learners (Fung & Carter, 2007), Belgium English learners (Buysee, 2012) or Turkish EFL learners (Asik &Cephe, 2013). The results showed differences in using EFL learners with different backgrounds. Aijmer (2004) pointed out that more experimental research should be conducted with different English learners in the world. Bolden (2015) also argued that the experimental studies nowadays were not

enough to have the whole picture on the use of discourse markers from the non-native speakers' point of view. Der (2010) argued that it should be encouraged to conduct research with different backgrounds or different criteria of EFL learners to examine deeper into the use of discourse markers in more detail. More data should be examined to see if they would be aligned with the previous research. Moreover, it was argued by Bolden (2015) that the research could not be so persuasive as the data or the statistics were not enough to examine the significant difference. Thus, this research tried to present statistics in detail to examine Thai English learners' use of spoken discourse markers compared to native speakers.

In the previous studies on the comparison between native English speakers and non-native English speakers, researchers studied different EFL learners at different English levels (Buysse, 2012; Diskin, 2017; Fung & Carter, 2007; House, 2013; Polat, 2011), i.e., the novice of EFL learners (Hellermann & Vergun, 2007); Intermediate-level (B1-level) EFL learners (Bax, Nakatsuhara & Waller, 2019; Sitthirak, 2013); advanced-level EFL learners (Aull & Lancaster, 2014; Sadeghi & Heidaryan, 2012). Based on Cambridge English Assessment of International language standards (2001), this research was intended to investigate B1-level and C1-level English learners of Thai Nationality to study how both levels use spoken discourse markers in conversation compared to native English speakers.

Some previous research studied English written discourse markers used by Thai EFL learners (Jangarun & Luksaneeyanawin, 2016; Prommas & Sinwongsuwat, 2011; Sitthirak, 2013). Prommas and Sinwongsuwat (2011) and Jangarun and Luksaneeyanawin (2016) studied discourse connectors in argumentative essays written by Thai EFL learners. Sitthirak (2013) explored contrastive discourse markers and used test-form assessment to examine whether Thai EFL learners were able to identify contrastive discourse markers. In contrast, Arya (2020) and Nookam (2010)

5

investigated the spoken discourse markers used by Thai EFL learners. However, it is overtly to see that the studies of spoken discourse markers used by Thai EFL learners were not enough. Few studies were found to discover how Thai EFL learners used the English spoken discourse markers in English conversation and the differences in using them compared to native English speakers.

Under the circumstances, this research was intended to fill the gaps as follows.

First, this research examined the use of English spoken discourse markers by Thai EFL learners in English conversation from two perspectives: the frequency and the pragmatic functions of the spoken discourse markers.

Second, this research compared the differences in both the frequencies and the pragmatic functions of the English spoken discourse markers used by Thai EFL learners and native English speakers, wherein the problems of using the English spoken discourse markers by Thai EFL learners were found at the same time.

1.2 Research questions

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

2. Is there a significant difference in the use of spoken discourse markers between Thai B1-level EFL learners and Thai C1-level EFL learners?

3. What are the pragmatic functions of spoken discourse markers used by Thai EFL learners?

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

1.3 Objectives of the study

1. To explore whether there is a significant difference in the use of spoken discourse markers between Thai EFL learners and native English speakers.

2. To explore whether there is a significant difference in the use of spoken discourse markers between Thai B1-level EFL learners and Thai C1-level EFL learners.

3. To identify the pragmatic functions of discourse markers used by both Thai B1-level EFL learners and Thai C1-level EFL learners.

4. To identify the different pragmatic functions of discourse markers used by Thai EFL learners compared to native English speakers.

1.4 Statement of hypothesis

1. There is a significant difference in the frequency in the use of spoken discourse markers between Thai EFL learners (including both B1-level Thai EFL learners and C1-level Thai EFL learners) and native English speakers.

2. There is a significant difference in the frequency in the use of spoken discourse markers between B1-level Thai EFL learners and C1-level Thai EFL learners. Generally speaking, B1-level Thai EFL learners use spoken discourse markers at a lower frequency compared to C1-level learners.

3. The number of pragmatic functions found in using spoken discourse markers is in this order: B1-level Thai EFL learners < C1-level Thai EFL learners < native speakers.

4. Both-level Thai EFL learners use basic pragmatic functions of spoken discourse markers and mainly tend to stick to using certain pragmatic functions repeatedly.

1.5 Scope of the study

1. Previous research focused on discourse markers in both spoken form and written form (Fraser, 1988, 1993; Modhish, 2012; Rahimi, 2011). However, the discourse markers in these two forms were not the same (Bolden, 2015). It is not possible for this research to focus on discourse markers in both forms. Since Aijmer (2004) pointed out that discourse markers should be focused more on the ones used in oral communication, especially in daily conversations (Fung & Carter, 2007; Haselow, 2011; Prabhumoye, 2017; Tagliamonte, 2005; Trillo, 2002). Hence, this research only focuses on spoken discourse markers in conversations.

2. Although there has not been an agreement in the inventory of discourse markers in previous studies (Borderia, 2006; Fraser, 2006, 2011), certain discourse markers used by native English speakers have always been focused and studied, including *well* (Lam, 2009; Popescu-Belis & Zufferey, 2010; Waltereit & Detges, 2007), *you know* (Fitzmaurice, 2004; Fung & Carter, 2007; Polat, 2011; Trillo, 2002; Vanda & Peter, 2011), *like* (Polat, 2011; Popescu-Belis & Zufferey, 2010), etc. In the previous studies, six English spoken discourse markers used by native English speakers at high frequencies in various pragmatic functions were commonly investigated (Aijmer, 1997, 2011; Brinton, 2008, 2010). Under the circumstances, these six English spoken discourse markers were studied in this research for investigating how Thai EFL learners used them in English conversation. They are *like*, *so*, *well*, *you know*, *I think* and *I mean*.

3. The population and the participants in this research focused on Thai undergraduates, whose English levels should be above B1 level based on the CEFR standard (Council of Europe, 2001). In order to examine the significant difference at different English levels of Thai EFL learners, this research focused on the participants in Thai B1-level EFL learners as the intermediate-English-level EFL learners and

Thai C1-level EFL learners as the advanced-English-level EFL learners. Meanwhile, this research used the CU-TEP score for examining the English level of each participant. Moreover, Thai EFL learners in this research refer to the L1 Thai EFL learners who were raised and lived in Thailand to the undergraduate level.

1.6 Definition of terms

   1.6.1 Discourse markers

This research mainly uses the definition proposed by Brinton (1996, 2008). This is because Brinton illustrated the development of the pragmatic functions of spoken discourse markers from their original semantic sense by using some specific spoken discourse markers, i.e. *I mean*, and summarized the concept of spoken discourse markers comprehensively from their characteristics to the framework of their functions by concluding and developing ideas from other researchers, such as Schiffrin (1987), Fraser (1988, 1999), etc. Brinton's definition has also been used by other researchers (Alami, 2015).

First, to give a short definition of spoken discourse markers, they are defined as "phonologically short items that have no or little referential meaning but serve pragmatic or procedural purpose" (Brinton, 1996, 2008).

A discourse marker is supposed to be a short lexical item in terms of its phonology. A discourse marker basically has no referential meaning. What makes a short lexical item to be a discourse marker is that it possesses procedural meaning and serves itself in different pragmatic functions in utterances.

The pragmatic functions of a discourse marker should be negotiated by the context (Fraser, 1999). It signals a relationship between the upcoming discourse or utterance and the prior discourse or utterance. The meanings and the relationship are both based on and concerned with speakers' communicative intention, where their

pragmatic functions make an effect on hearers to inform hearers how to interpret speakers' discourses or utterances. This intrinsic feature (Brinton, 2008) is shown by discourse markers' pragmatic functions. Their functions guide hearers on how to interpret the orientation of the speaker's point of view.

1.6.2 Thai EFL learners

The population and the participants in this research refer to L1 Thai-nationality EFL learners who were raised and lived in Thailand to the undergraduate level. This research included L1 Thai B1-level and L1 Thai C1-level EFL learners based on the CEFR assessment standard (Council of Europe, 2001). According to the CEFR standard, if IELTS scores are between 4.0 and 5.0, they will be marked as B1 level. If IELTS scores are between 6.5 and 8, they will be marked as C1 level.

Based on Wudthayagorn (2018), the CU-TEP scores have mapped with the scores of CEFR. According to the research, CU-TEP has the cut-off scores range with the equality of CEFR levels from A2 to C1, including B1 and C1 levels. CU-TEP total scores are 120 points. The cut-off points of B1 level are 35-69 points. The cut-off points of C1 level are 99-120 points. Hence, the CU-TEP scores were considered as the instrument for choosing the participants in both B1-level and C1-level.

1.6.3 Pragmatic functions of spoken discourse markers

Based on Brinton's research (1996, 2008, 2010), different discourse markers had different pragmatic functions in communication. However, they can be concluded into two major functions: textual functions and interpersonal functions. In textual functions, discourse markers can mark boundaries or mark turn-taking in communication. In interpersonal functions, discourse markers are used either from the subjective point of view or from the interactive point of view. In this research, the pragmatic functions of spoken discourse markers are separated into textual functions and interpersonal functions based on Brinton's (2008) framework. Each discourse

marker under investigation has its unique functions. They will be analyzed accordingly.

1.6.4 Learner corpora

McEnery et al. (2006) proposed that learner corpora should be "collection of machine-readable authentic texts (including transcripts of spoken data) which is sampled to be representative of learners' interlanguage" (p. 5). What makes learner corpora special is that the original data is collected from the language learners. In many researchers' eyes, their language can be called inter-language, which is worth studying. The representativeness of the data in the learner corpora can be treated as the language patterns of the language learners themselves. Thus, learner corpora may be seen as the huge collection of original language data directly from language learners.

1.7 Significance of the study

This study seeks to find out how Thai EFL learners produce spoken discourse markers in oral communication. It will benefit linguists to understand differences in using the spoken discourse markers between two different-level Thai EFL learners.

This study seeks to investigate the pragmatic functions of spoken discourse markers used by Thai EFL learners and compare them to native English speakers. The findings will reveal differences in using the pragmatic functions by having the comparisons between Thai EFL learners and native English speakers.

This research fills the gap in how Thai EFL learners use spoken discourse markers in English conversation. The results will present problems with the use of the spoken discourse markers by Thai EFL learners in English conversation. It is useful for future English teaching and curriculum design in Thailand. Thai EFL learners can

also discover their errors in using spoken discourse markers and try to pay more attention to them in future English studies.

The findings of this research may provide more information on how non-native English speakers use spoken discourse markers in communication. The results can attest to whether they are aligned with the result of the previous research to examine the patterns of how EFL learners in different backgrounds use spoken discourse markers.

Chapter 2

Literature review

The literature review is presented in two sections: previous studies on discourse markers and previous studies on learner corpora.

In the first section, based on the previous studies on discourse markers from different perspectives, five main parts are presented. In the first part, basic issues of discourse markers are reviewed (Aijmer, 1997; Fraser, 1988; Levinson, 1983). The basic issues of discourse markers include definition, features, classification, and basic functions. In the second part, previous studies on discourse markers in written data are reviewed in general. These studies are mainly experimental studies that explore non-native speakers' use of discourse markers in English writing (Aull, 2019; Babanoglu, 2014; Modhish, 2012). In the third part, previous studies on discourse markers in second language acquisition are reviewed in general. These studies focus on how EFL learners acquire discourse markers and suitable teaching methods for learners to learn discourse markers from four language skills perspectives (Alghamdi, 2014; Jones & Carter, 2012; Pasaribu, 2017). In the fourth part, previous studies on discourse markers in spoken data are reviewed. This part is the main part of the literature review because it is directly related to the topic of this research. Previous studies on native speakers' use of discourse markers (Norrick, 2011; Tagliamonte, 2005), studies on non-native speakers' use of discourse markers (Hellermann & Vergun, 2007; Vanda & Peter, 2011), comparative studies (Aijmer, 2004; Trillo, 2002) and previous studies on pragmatic functions of the six discourse markers that are investigated in this research are fully reviewed. In the fifth part, previous studies on English discourse markers in Thailand are reviewed in general.

In the second section, previous research related to learner corpora is reviewed and presented in four parts: design of learner corpora (Granger, 1998a; Tono, 2003),

13

learner corpora with experimental studies (Callies, 2015; Neff et al., 2004), technical issues for dealing with learner corpora (Passarotti et al., 2009; Ramos et al., 2010), and studies on learner corpora and second language acquisition (Granger, 2013, 2018).

This research starts with a literature review of previous studies on discourse markers first.

2.1 Basic issues on discourse markers

Most of the basic issues on discourse markers were studied from the 1980s until the end of the 20th century. Four areas of basic issues on discourse markers have been concerned by researchers, which were the issue of terminology, the issue of defining discourse markers, the main features of discourse markers, the categorization of discourse markers and the basic functions of discourse markers. Researchers, such as (Brinton, 1990), Fraser (1988, 1993, 1996, 1999), mainly used written data from native English speakers to explore the issues mentioned above at an early stage. Meanwhile, researchers, such as Schiffrin (1987), Brinton (1996), Aijmer (1997), mainly used spoken data to study discourse markers. The results from these early studies became the foundation for future studies on non-native English speakers' use of discourse markers.

2.1.1 Terminology and definition of discourse markers

It was found that various perspectives contributed to the terminology of the discourse markers, followed by the definition of the discourse markers.

Levinson (1983) used "words and phrases" (p. 87) to refer to discourse markers. After Levinson, different researchers used different terms to refer to discourse markers. Schourup (1985) presented the term "discourse particles" in his early work. The examples he presented in the research were like, well and you know, which were studied more in research afterward. However, Schourup (1999) changed the term

"discourse particles" into "discourse markers". What he argued was that the term "particle" should be in syntactic term from the historical perspective. Schiffrin (1987) used the term "discourse markers" in her work. Since Schiffrin (1987), whose work influenced hugely to the research afterward, the term "discourse markers" started to be accepted and used by other researchers. Schiffrin (1987) defined discourse markers as "sequentially dependent elements which bracket units of talk" (p. 31). The definition firstly claimed that discourse markers were dependent elements. They appeared in between discourses. Until 1999, researchers had found that discourse markers could not be simply categorized into syntactic lexical items. Since then, the term "discourse markers" was used the most by different researchers (Brinton, 1990; Fraser, 1990, 1993, 1996, 1999, 2006, 2011, 2015; Heeman et al., 1998; Jucker & Ziv, 1998). Most of the recent studies used this term (Blakemore & Gallai, 2014; Dobrovoljc, 2017; Polat, 2011; Sakita, 2013).

The issue of defining discourse markers was at the same time as the issue of terminology. In the previous research, how discourse markers were defined can be seen from two perspectives. It should be noted that these two perspectives were not utterly independent of each other but completed the cognition of discourse markers comprehensively.

The first perspective was mainly focused on the textual functions of discourse markers. Levinson (1983) argued that some "words and phrases" in English indicated the relationship between an utterance and the prior discourse. Examples of these "words and phrases" that Levinson presented were but, however, anyway, well, actually, so, after all, etc. He pointed out that these "words and phrases" did not influence the truth-conditional treatment but indicated in a complex way in which how the utterances that followed these "words and phrases" responded to the prior utterance. Levinson did not demonstrate what the relationships could be in detail, but

his argument illustrated clearly that these "words and phrases" had certain textual functions that indicated a certain relationship between the prior utterance and the following utterance.

Blakemore (1987) used Relevance Theory (Sperber & Wilson, 1986) to study discourse markers and also treated them at the textual level. He argued that by using discourse markers, the relevance of one discourse segment to another would be dependent, and the discourse markers would be the expressions that imposed "constraints on relevance in virtue of the inferential connections they express" (p. 105).

Another example was from Fraser (1988), arguing that discourse markers signaled a relationship between the prior discourse (S1) and the following discourse (S2). They were "a pragmatic class, lexical expressions drawn from the syntactic classes of conjunctions, adverbials, and prepositional phrases" (p. 950). It can be seen that Fraser treated the textual functions of discourse markers significantly.

Similar definitions that focused on the textual functions of discourse markers can be found mainly in the 1980s and 1990s. Goldberg (1980) argued that they were marking devices that displayed "relevance to the information set as established by the immediately preceding contribution" (p. 141). Redeker (1990) believed that they were the linguistic expressions that were used to "signal the relation of utterance to the immediate context" (p. 372).

Soon enough, discourse markers were discovered that not only had textual functions but indicated the relationship between the speaker and the hearer, which led to the second perspective.

In contrast to the discourse markers' textual functions, the second perspective paid more attention to their interpersonal functions. The emergence of this perspective was aligned with the phenomenon that more research started to focus on the discourse

markers in oral communication rather than the ones in the written form (Aijmer, 1997; Bazzanella, 2006).

Stenström (1995) argued that "they (discourse markers) were highly neutralized in meaning, and they were person-to-person-oriented and socially required" (p. 299). The definition reflects that the functions of discourse markers are more concerned with their roles in social interaction from the people-oriented perspective. Based on the analysis of the spoken data, Stenström (1995) found out that the discourse markers were syntactically "deletable", and pragmatically required in discourse. Hence, besides textual functions, discourse markers were found that they had interpersonal functions in communication.

More definitions from this perspective can be found after the 21st century. 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 includes two pieces of information. For one thing, the discourse markers are used to show the attitude of the speaker; for another, the speaker shows the attitudes for a significant reason that the speaker intends to interact with the hearer. Bazzanella (2006) proposed the definition from the same perspective 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 was aligned with the findings of plenty of research (Aijmer, 2016; Fung & Carter, 2007; Sakita, 2013). For example, in Sakita's research (2013) on the discourse marker well, Sakita argued that well-worked "as a resource for management of the relationship between various stances" (p. 82). Speakers are faced with the need to manage relations among stances as a way of dealing with "socio-cognitive relations" (Du Bois, 2007).

It should be noted that some researchers illustrated the definition in a relatively short sentence. But in fact, they provided a more additional explanation.

Schiffrin (1987) defined discourse markers as "sequentially dependent elements which bracket units of talk" (p. 31). It seems that she also paid more attention to the textual functions of discourse markers. But with her framework of functions of discourse markers, she proposed two structures that related to show the relationship between the speaker and hearer.

Schiffrin's framework of functions of discourse markers is also known as the five planes of talk (1987, p.24-28). She believed that the conversation was a multilayered interaction, and discourse markers' functions would be seen in the pragmatic layer as in five planes of talk as follows.

1. Exchange structure (ES), which reflects the mechanics of the conversational interchange and shows the results of the participants' turn-taking and how these alternations are related to each other.

2. Action structure (AS), which reflects the sequence of speech acts that occur within the discourse.

3. Ideational structure (IdS), which reflects a certain relationship between the ideas (propositions) found within the discourse.

4. Participation framework (PF), which refers to the different ways in which speaker and hearer can relate to each other.

5. Information state (InS), which reflects the ongoing organization and management of knowledge as it evolves over the course of the discourse.

In this framework, the exchange structure and participation framework focus on the relationship between the speaker and the hearer, while the rest focus on the relationship at the textual level.

Following the ideas of the researchers mentioned above, Brinton (2008) proposed a definition for discourse markers that were used in this research. At the heart of the definition by Brinton, she proposed that discourse markers were "phonologically short items that have no or little referential meaning but serve pragmatic or procedural purpose" (2008, p. 1). First, discourse markers were short items in phonology. Usually, they were lexicons or lexical items with two or three lexicons. Second, what served discourse markers as discourse markers was that they should have pragmatic or procedural meanings during the communication other than their referential meaning in the context.

In conclusion, when discourse markers were defined in the previous research, the essence of the discourse markers that researchers have laid emphasis on was their pragmatic functions in discourse. Both textual functions and interpersonal functions have been the key points for researchers to define discourse markers. Moreover, some researchers also involved the syntactic features, phonological features, lexical features, or other features of discourse markers in the definition. The discussion of the features of discourse markers will be presented in the following part.

2.1.2 Main features of discourse markers

The study of the features of discourse markers can be traced back to Fraser's (1988) research. Other researchers also tried to categorize them from their different perspectives. Eight issues of categorization and features of discourse markers should be mentioned based on previous studies.

First of all, based on many studies from the 1980s until the 2010s, it was illustrated that discourse markers were one single word, such as *so*, *well*, *now*, and also were a combination of two to four words, such as *I mean*, *you know*, *on the other hand* (Fraser, 1988, 1990; Aijmer, 1997, 2004; Sakita, 2013; Schourup, 1985). Recently, some researchers also started to study combining discourse markers (Fraser,

2011). These combining discourse markers were a combination of two or more discourse markers, such as *oh anyway*, *but anyway*.

Second, as has been shown above, discourse markers came from different part-of-speech if they were one single word. The combining discourse markers may come from different part-of-speech words.

Third, from the perspective of phonology, they were usually unstressed or reduced (Bolden, 2015; Brinton, 1990; Jucker & Ziv, 1998). However, in recent studies, researchers also found that not it was not always the same in each context (Dehe & Wichmann, 2010). Their stress or reduction in the phonological point of view should depend on specific situations.

Fourth, the position of discourse markers was usually at the initial of a sentence or an utterance (Aijmer, 2016; Fraser, 1993, 1996; Jucker & Ziv, 1998; Schourup, 1985). However, recent studies also found that there would be situations where speakers put the same discourse markers into different positions of utterances (Diskin, 2017; Haselow, 2011; Lam, 2009).

Fifth, discourse markers were usually treated as syntactic independence (Brinton, 1990, 1996; Fraser, 1988), indicating that even if they were deleted, their absence should not affect the semantic meaning of the utterance or the sentence (Bolden, 2015; Fraser, 1988). From this point of view, it was also suggested that what made discourse markers important was their pragmatic functions, rather than their semantic meaning.

Sixth, researchers pointed out that discourse markers did not have propositional meaning, but had procedural meaning (Blakemore, 1996, 2002; Borderia, 2008; Der, 2010; Fraser, 1996, 1999; Wilson & Sperber, 1993). They were not activated to be one part of the members that were responsible for expressing the semantic meaning to hearers or readers. What they only have was the procedural meaning, which came from their pragmatic functions.

Seventh, since Schourup (1985) used interview conversations as original data to study discourse markers, some researchers believed that discourse markers were a feature of spoken language (Aijmer, 2004). In general, it was still believed that discourse markers appeared both in spoken form and written form. However, through studies on discourse markers in both forms, it was overtly to see that discourse markers in spoken form were different from those in written form, such as the use of well, you know, anyway, which were generally not used in written form, especially in the argumentative essay (Aijmer, 1997; Landgrebe, 2012; Tree, 2015). It was argued that some discourse markers belonged to the spoken language, which interested more researchers to study them in oral communication (Aijmer, 2011).

Eighth, since Levinson (1983), Fraser (1988, 1990, 1993, 1996), Aijmer (1997, 2016), Schourup (1985, 2011), until all the recent empirical studies on discourse markers (Fung & Carter, 2007; Haselow, 2011; Prabhumoye et al., 2017; Tagliamonte, 2005; Trillo, 2002), discourse markers were found to be multi-functional. Discourse markers usually have more than one function in contexts. Different discourse markers in different genres of language also may perform different functions. Each discourse marker studied in previous research had more than one function in utterances. Therefore, multi-functionality has also been proposed by different researchers (Aijmer & Simon-Vandenbergen, 2011; Fraser, 1996).

2.1.3 Categorization of discourse markers

The categorization of discourse markers needs to be discussed from two points: the categorization of discourse markers as a whole range and the categorization of discourse markers within the range.

For the whole range of discourse markers, researchers categorized them into "the pragmatic class" (Aijmer, 2011, 2016; Brinton, 2008, 2010; Fraser, 1999). It is simply for the reason that they mainly serve pragmatic functions. Besides, Schourup (1999)

also argued that this "pragmatic class" was an open class, for there was no agreement on the list of discourse markers due to different perspectives of the researchers and the development of discourse markers that had been influenced by the language change based on Brinton's diachronic research on discourse markers (1996). The categorization of the whole range of discourse markers seems to have no or little agreement.

Some other perspectives of categorizing the whole range of discourse markers suggested categorizing them with their main features. Even though some features are not universal for all discourse markers, some common features can help categorize the whole range of discourse markers. For example, syntactically speaking, they are independent short items; and with the feature of no or little propositional meaning, they do not influence the semantic meaning of the discourse.

To categorize discourse markers within the range, they can be categorized by their different features.

In terms of lexical features, discourse markers are categorized by length of the discourse markers. Most of the recent research believed that discourse markers were one single word, such as *well*, *so*, *but*, *and*, *like*, *anyway*, *now*, etc.; or they were multi lexical items, such as *I mean*, *you know*, *you see*, *I think*, etc. (Aijmer, 2011).

In terms of the syntactic features, some researchers argued that discourse markers usually appeared in the initial position of an utterance (Aijmer, 2016; Fraser, 1993, 1996; Jucker & Ziv, 1998; Schourup, 1985), such as the example of using *well* as a discourse marker below.

A: What am I going to do now?

B: *Well*…I really don't know. (Fraser, 1999, p. 942)

However, not all the discourse markers can be categorized into the utterance-initial position. Some discourse markers were argued that could appear in the middle

of the utterance, such as *like*. It may appear in the middle of an utterance in the example shown below.

David: But yeah I've been doing *like* language exchanges and stuff.

(Diskin, 2017, p. 154)

Some researchers also categorized discourse markers from the perspective of their appearance in the written form and the spoken form. It was argued that some discourse markers such as *well*, *you know*, *I mean* were mainly used in spoken English (Aijmer, 2016). Thus, these discourse markers should be categorized into spoken discourse markers.

Schiffrin (1987) and Fraser (1999) categorized discourse markers by their main functions that served in discourse. Schiffrin (1987) studied eleven discourse markers and categorized each discourse marker under its main function as shown in Table 1.

Table 1: Categorization of the eleven discourse markers in planes of talk (Schiffrin, 1987, p.316)

Exchange structure	Action structure	Ideational structure	Participation framework	Information State
you know		*you know*	*you know*	*you know*
or		*or*		
and	*and*	*and*		
so	*so*	*so*	*so*	*so*
	because	*because*		*because*
	then	*then*		
	oh		*oh*	*oh*
		I mean	*I mean*	*I mean*
		now	*now*	
well	*well*	*well*	*well*	*well*
but	*but*	*but*		

Although each discourse marker under investigation had more than one plane of talk, Schiffrin believed that each discourse marker only had one main function. For example, the following discourse markers *but, and, or, so, because, now* and *then* had the main function of the ideational structure. The discourse markers *well* and *I mean* were mainly in charge of the expression or the negotiation of the relations between the speaker and the hearers.

Fraser (1999) proposed four types of discourse markers based on different kinds of relationships between the prior utterance and the following utterance. They were the contrastive markers (such as *but, yet, instead,* etc.); the elaborative markers (such as *in fact, moreover, and,* etc.); the inferential markers (such as *therefore, thus, after all,* etc.); the topic-change markers (such as *by the way, before I forgot,* etc.) However, Fraser categorized discourse markers only by their functions in written English.

In conclusion, discourse markers in the whole range were categorized as "an open pragmatic class" based on the previous research. Within the range of discourse markers, they were also categorized mainly from their different functions or different features. However, due to the fact that many recent studies only focused on an individual discourse marker, it should be noted that it could not simply identify which function would be the main function of a discourse marker, for the same discourse marker might be studied in several studies but had different results due to different perspectives, different resources of the data, different methodologies, etc. Thus, recent research seldom focused on the categorization within the range of discourse markers.

2.1.4 Basic functions of discourse markers

In the same situation as the definition of the discourse markers, the functions of discourse markers have not been in total agreement. There have been no strict rules to follow so far. In early research, researchers tried to conclude the functions of

discourse markers with their own perspectives (Fraser, 1988, 1990, 1993, 1996, 1999, 2015; Schiffrin, 1987). Thus, the conclusions from different research were not unanimous. In recent empirical research, researchers tried to study of specific discourse markers to discover the functions of each specific discourse marker.

Jucker and Ziv (1998) examined the functions of written discourse markers. They found that in most situations, the written discourse markers were generally in the category of discourse connectors, for their major use in academic writing was to connect utterances, such as *on the one hand*, *on the other hand*, etc.

Fraser (1988, 1993) used written English of native English speakers to investigate the basic functions of the discourse markers. Three types of the discourse markers were found. The first type was topic markers. Under this type, there were two subgroups: signaling the shift to a new topic and signaling a refocusing on the recurrent topic. The second type was discourse activity markers. These activities referred to types of discourse such as explaining or summarizing, but not to the type of message the speaker conveyed through the utterance (p. 28). Activity markers signaled that the current discourse activity was relevant to the foregoing discourse. Classifications in this type were plenty, such as explaining, interrupting, repeating, sequencing, summarizing, etc. The third type was message relationship markers. They signaled the basic relationship between the foregoing discourse and the upcoming discourse. There were four subgroups of this marker: parallel discourse markers; contrasting discourse markers; elaborative discourse markers and inferential discourse markers.

Fraser (1999) presented two points related to the functions of discourse markers. First, discourse markers were supposed to be one kind of lexical classification that linked some discourses. They should be in the middle of one foregoing utterance and one upcoming utterance. Thus, the model should be "S1+DM(s)+S2". This mode also

25

presented a basic function of discourse markers. They were connectors of two utterances. However, the research also mentioned that discourse markers, especially in the conversations, were not always directly linking the exact foregoing utterance to the upcoming one. In the conversations, people might say sentences in segments and details were scattered in several utterances. Second, Fraser tried to exclude some markers and explained again why those markers were not discourse markers. For example, even though *frankly* would be a word to express the speaker's attitude towards one thing, it did not have the function of linking the foregoing utterance. From this point of view, discourse markers in the written data had the function of linking discourses, which is aligned with the first point.

Fraser's (1999) research concluded the basic functions of discourse markers in the written data used by native English speakers. Besides Fraser, Aijmer (1997) used one single discourse marker *I think* to study its functions in discourses. It was one of the early research that used one single case to study the functions of discourse markers.

Aijmer (1997) studied the spoken discourse marker *I think*. Based on the research, *I think* could be categorized as three functions from the pragmatic perspective. First, it was similar to *I believe*, which indicated that the speaker was supposed to really know or understand something so that the speaker used *I think* to express what he was about to say was believable in that his attitude was as certain as it was. However, this situation was similar to the difference between *may* and *might*. The hearer would think that compared to the phrase *I believe*, *I think* was slightly into the balance of uncertainty in a more or less kind of way. Thus, it came to the second function that *I think* was used as an expression in terms of its pragmatic function to prove that the speaker was quite certain about what he or she was about to say and made a sign that the speaker was personally sure about the utterance nearly 100%

guarantee. However, the research also found that in many situations, speakers added this phrase at the beginning of the utterance to show that he or she might still have doubts or uncertainty about the following utterance. It was therefore in a tentative function to express the uncertainty. The research continued its study on the combination of *I think* with other discourse particles, such as *well I think* or *I think I mean*, etc. In conclusion, Aijmer (1997) presented a whole picture of the spoken discourse marker *I think*, in which *I think* was separated into the traditional modality and made the new term "discourse marker" with its unique pragmatic function introduced to future studies.

Based on Brinton (1996, 2008), the basic functions of discourse markers were separated into interpersonal functions and textual functions. In interpersonal functions, there were two domains. The first domain was that speakers used discourse markers to subjectively express their personal attitudes towards something or someone. The second domain was that discourse markers were used to interactively achieve intimacy between speakers and hearers. It was thus conceivable that, in interpersonal functions, discourse markers should be regarded as a signal to maintain the relationship between the speaker and the hearer. Some research suggested that to show the relationship between the speaker and his/her orientation towards the produced discourse was considered an intrinsic feature of discourse markers (Alami, 2015). In textual functions, there were also two domains. The first domain was to mark different types of boundaries among discourses. The second domain was to be the signal for turn-taking. In textual functions, various functions were found in detail, such as discourse markers as fillers, topic shifting or switching, turn-taking, a signal for opening or closing a discourse, information indicators and sequence or relevance markers, etc.

It is worth mentioning here that the review above on Brinton's (2008) introduction of the basic functions of discourse markers has not been into more concrete details. Brinton's (2008) framework of functions of discourse markers will be presented in the following part in detail.

To sum up, researchers who studied discourse markers from different perspectives may achieve different functions (Yilmaz, 2004). Studies on the functions of discourse markers have kept continuing since no agreement has been illustrated (Bolden, 2015). However, the previous research on functions of discourse markers has provided a strong foundation for the further studies to explore more functions of discourse markers in different contextual environments.

2.2 Studies on written discourse markers

Based on the review of the basic issues of discourse markers in the previous part, some early studies in the last century used native speakers' written data to investigate the use of written discourse markers (Brinton, 1990, 1996; Fraser, 1988, 1993, 1996, 1999). Thus, researchers intended to further their studies on how non-native English speakers used discourse markers in written form.

Under the circumstances, two categories of studies on discourse markers in written form have been carried out: studies only focused on non-native English speakers' inappropriate use of discourse markers in English writing (Li, 2010; Modhish, 2012; Rahimi, 2011); studies that compared the different uses of discourse markers between non-native English speakers and native English speakers, or compared the different uses of discourse markers among different genres of non-native English speakers' English writing. (Aull, 2019; Aull & Lancaster, 2014; Babanoglu, 2014).

2.2.1 Studies on written discourse markers used by non-native English speakers

As has pointed out above, some research only focused on studying the written discourse markers used by non-native English speakers (Li, 2010; Modhish, 2012; Rahima, 2011). They mainly explored the pattern in the use of written discourse markers by non-native English speakers. Two issues interested researchers in general. For one thing, researchers explored what written discourse markers non-native English speakers tended to use in English essays. For another, how non-native English speakers used them in English essays.

Li (2010) investigated the written discourse markers used by Chinese EFL learners in English writing. The research tried to find out the misuse and the inappropriate use of written discourse markers by Chinese EFL learners. The findings suggested that a good use of written discourse markers was one of the main conditions for good English writing. It could be seen that without written discourse markers, students' writing would not be coherent and no logical explanation. It was found that for the misuse of discourse markers, students had different problems in every aspect, such as grammatical errors in structure. By a closer investigation of some essays, it was found that some Chinese EFL students did not understand the structure of the written discourse markers among sentences. They could use them in a useless structure. Moreover, another problem regarding the inappropriate use of the written discourse markers was the pragmatic functions. It seemed that some Chinese EFL learners did not use some functions correctly in different contexts.

Rahimi (2011) investigated the written discourse markers in two different genres of academic essays: argumentative essays and expository essays. The research was intended to discover the frequency and the type of discourse markers used in the two genres. The study used 56 Iranian English major students as the participants. They

were asked to write two essays in two given genres. The results showed that the participants used different written discourse markers in two genres of essays. However, it was found that the use of the written discourse markers in argumentative and expository essays did not significantly predict the quality of the essays, indicating that the use of the written discourse markers may not influence the quality of the essays in both genres.

Modhish (2012) investigated the use of the written discourse markers by Arab EFL learners in their English writing. It was intended to find whether Arab EFL learners acquired discourse markers *well* in their English writing, and how they used discourse markers in the writing. The research also had a similar result as the one in Rahimi (2011) that Arab EFL learners used almost all types of written discourse markers, including contrastive markers, causative markers, topic relating markers and elaborative markers. This result generally showed that learners had an awareness of using the written discourse markers in English writing and they somehow knew the situations where they should use them in different contexts. However, the research also found that there was no strong positive correlation between the quality of learners' essays and the use of the written discourse markers. Most of the learners still used other lexical items to analyze topics, which was commonly based on the data. Thus, with the limited length of the essays, learners had no more room for some written discourse markers. In the comparisons with other lexical items, learners used fewer written discourse markers. The top-used written discourse markers were the common ones, such as *and*, *also* or *so*. Meanwhile, learners tended to combine some written discourse markers together, i.e., *and also*.

Khedri et al. (2013) investigated interactive discourse markers in academic research articles published in academic journals in two disciplines. They intended to figure out how writers chose discourse markers in abstract to interact with readers and

30

how discourse markers worked in this situation. The study discovered that transitional markers had their leading position in the abstracts. Meanwhile, in different types of abstracts, researchers used different discourse markers to illustrate their attitudes. It was illustrated by the study that researchers in different areas tended to have their own ways to choose the interactive discourse markers for their writing styles.

Asr and Demberg (2013) investigated written discourse connectives. The research was intended to study the written discourse connectives' roles in texts, explored their functions and tried to examine their relations with the coherence of the texts. With the markedness measurement, the research categorized data and found four groups of discourse connectives related to the coherence of texts. They were temporal markers, contingency, expansion and comparison. The research tried to present the hierarchy levels of discourse connectives and related them in all dimensions to examine them from more perspectives.

Hence, the findings in different studies above indicated that discourse markers in written form somehow influenced EFL learners to write English essays. In many instances, EFL learners had the ability to use the written discourse markers in English writing. However, some EFL learners had a problem with the inappropriate use of certain written discourse markers. For one thing, EFL learners had limited knowledge of the English written discourse markers; for another, EFL learners were lack of fully understanding of certain functions of some written discourse markers so that they might use certain written discourse markers in a wrong way.

2.2.2 Comparative studies on written discourse markers

Two types of comparative studies on discourse markers have been discovered. First, a number of studies made comparisons between the use of the written discourse markers between native English speakers and non-native English speakers (Alghamdi, 2014; Povolna, 2012). Second, some studies compared different groups of participants

with different backgrounds in the same genre of essay to examine the different uses (Aull, 2019; Pasaribu, 2017).

Povolna (2012) explored the use of the written discourse markers by novice non-native English writers in academic writing. The research was intended to investigate which written discourse markers novice students would choose when they wanted to express casual and contrastive relations and whether they could choose the right discourse markers correctly. The results illustrated that for casual discourse markers, the ones that novices used were almost the same. However, the use of contrastive discourse markers was uneven and had big differences. Different written discourse markers in both types would be placed in different positions in utterances. One single discourse marker could be placed into the initial position of the sentence or in the middle of the sentence. The findings indicated that the situations of using casual and contrastive discourse markers were totally different. The functions of contrastive discourse markers seemed to be more interesting for novices to choose from. Based on the qualitative analysis, researchers also pointed out that there were situations where novices used discourse markers in the wrong way. Some novices tended to use only a small range of discourse markers in writing.

Alghamdi (2014) investigated the use of the written discourse markers in ESL personal narratives and argumentative essays in both non-native English speakers and native English speakers. It was found that non-native English speakers used more written discourse markers than native English speakers in both personal narratives and argumentative essays. However, in both types of writing, there was no significant difference in using the written discourse markers.

Babanoglu (2014) also investigated the use of pragmatic markers in EFL learners in argumentative essays. The study aimed at discovering whether there was a significant difference in using the pragmatic markers in argumentative essays between

native English speakers and EFL learners. It built two learner corpora and one native speaker corpus for comparison. The result showed that EFL learners used different written discourse markers in argumentative essays compared to the use of native English speakers. Meanwhile, EFL learners overused pragmatic markers of the written discourse marker in English writing compared to native speakers. Therefore, the study indicated that EFL learners should spend time specifically learning the pragmatic markers. They need to be given examples of how native speakers used these markers. The study also suggested that the phenomenon of overusing pragmatic markers should be paid much more attention in the future pedagogical implication.

Muşlu (2018) explored the written discourse markers as stance markers in the academic writing of Turkish EFL learners and made a comparison with native English speakers. It was intended to investigate the stance lexical bundles (SLB) in Turkish EFL learners' academic English writing with the discovery of the frequency list and the features of using SLB by Turkish EFL learners in academic writing. It was found that both Turkish EFL learners and native English speakers used the written discourse markers to point out personal stance. When using the certainty and uncertainty devices, it was found that both native English speakers and Turkish EFL learners used certainty devices more than uncertainty ones. Thus, the result indicated that native English speakers and Turkish EFL learners had a similar pattern of using the stance markers in argumentative essays.

Simeikaite (2012) had an interesting study on the spoken discourse markers in academic writing. The research explored the use of the spoken discourse markers in academic writing by both EFL learners and native English speakers. The findings illustrated that EFL learners tended to use more spoken discourse markers in writing compared to native English speakers.

To sum up, the previous research on the written discourse markers used by non-native English speakers compared to native English speakers indicated that EFL learners in different backgrounds had different uses of discourse markers in written form. In some situations, non-native English speakers may overuse the written discourse markers compared to native English speakers (Alghamdi, 2014; Babanoglu, 2014), whereas in some other situations, non-native English speakers may use the written discourse markers as similar to native English speakers (Muşlu, 2018; Povolna, 2012).

Besides the comparisons between native English speakers and non-native English speakers, some studies also focused on comparisons between different genres of writing written by non-native English speakers. These studies tried to discover differences in the use of the written discourse markers in different genres of writings by non-native English speakers, or differences in the use of the written discourse markers by non-native English speakers in different English levels.

Aull (2019) explored linguistic markers of stance used by EFL upper-level students in different genres of English academic writing. One genre was persuasive argumentative writing and the other one was analytic explanatory writing. The research used MICUSP to examine the use of these markers in upper-level students' English writing. It was concluded that upper-level EFL learners had a pattern of using linguistic markers of stance in both genres. Between the two genres, students used almost the same markers without a significant difference. These markers showed their significant roles in English writing, for they directly connected to the description of what students wanted to express. If they did not use them appropriately, it may influence their writing in both genres.

Bax et al. (2019) also investigated the written discourse markers used by EFL learners at two different levels of English writing. The research was intended to

explore how these two-level EFL learners used discourse markers in English writing and their differences in using them. The same to Aull (2019), it was found that the upper-level learners used more written discourse markers in the essays than intermediate-level learners in general. Both levels had a pattern of using different types of written discourse markers. Attitude markers, code glosses and logical connectives were among the high-frequency list used by both levels.

Pasaribu (2017) explored the use of the written discourse markers in the perspective of gender differences. The research was intended to find whether gender differences would also be one of the factors that influenced the different uses of discourse markers in writing academic essays. The result illustrated that male students used more written discourse markers than female students. Both male and female students followed the same pattern of using the written discourse markers. Meanwhile, both male and female students used similar discourse markers in each type. The findings indicated that gender differences might not be the major factor in influencing the use of the written discourse markers in argumentative essays.

Aull and Lancaster (2014) examined the use of the stance markers in academic writing between new college students and advanced students. The research aimed at examining whether new students understood linguistic markers of stance in English writing, how to use them and their use compared to advanced learners. The results showed that the first-year college students used hedges and boosters the most in all types of linguistic markers of stance. It was found that more advanced learners would use more hedges than boosters. Another type of linguistic markers of stance learners usually used was code glosses that "helped readers grasp the appropriate meanings of elements in texts" (Vande-Kopple, 1985).

Therefore, based on what has been reviewed, it can be concluded that the written discourse markers were used by different levels of EFL learners in English writing.

35

Learners had the awareness of using the written discourse markers in English writing. On the contrary, two main problems have also been discovered. On the one hand, EFL learners tended to only focus on a small number of the written discourse markers, such as *and, but, however,* etc. It has been argued that perhaps these written discourse markers were the ones that were most commonly taught by teachers so that learners might have a better awareness of using them in English writing (Aull, 2019; Povolna, 2012). It was assumed that EFL learners used a limited amount of the written discourse markers because they had not acquired enough written discourse markers. On the other hand, mastering all functions of each written discourse marker would be difficult for EFL learners.

2.3 Studies on discourse markers in second language acquisition

Since some studies have mentioned that non-native English speakers had the problems using the discourse markers compared to native English speakers (Khedri et al., 2013; Li, 2010; Modhish, 2012), researchers and teachers started to make students acquire discourse markers consciously through different teaching and learning procedures. Two types of research in this area have been found. The first research type was to examine different instructional approaches that could improve EFL learners to acquire discourse markers better (Hernandez & Rodriguez-Gonzalez, 2013; Jones & Carter, 2012). The second research type was to examine whether the improvement of acquiring discourse markers could improve EFL learners' basic English skills (House, 2013; Khatib & Safari, 2011). Since studies on discourse markers in second language acquisition are not at the core relevance of this research, they will be briefly reviewed.

Approaches to instruct discourse markers have been one of the main issues in this area. Hernandez and Rodriguez-Gonzalez (2013) did an experimental study on

the acquisition of L2 discourse markers in different instructional approaches. The research was intended to discover the different results by using different instructional approaches to teach discourse markers. It was found that there was a significant difference among three groups that had been taught discourse markers in three different methods. With the result of ANOVA, the research concluded that combined with explicit instruction, making students acquire discourse markers in a conscious way would help them improve their second language acquisition.

Jones and Carter (2012) also investigated different instruction approaches in teaching discourse markers for English speaking. The research tried to examine two approaches to explore which of them would be better to teach discourse markers in spoken English. The study found that PPP instructions would be the better instruction approach to teaching discourse markers in spoken English. With the survey of the learners, they also thought that PPP helped them understand much better on discourse markers and improve using them well.

Therefore, it can be seen that using proper instructions to teach discourse markers would help and improve EFL learners' use of discourse markers in English speaking. The two research also indicated that discourse markers were important for EFL learners to achieve a high level of speaking skill.

Another type of research mainly aimed at examining whether the acquisition of discourse markers would be significant to English proficiency and how discourse markers would influence four skills of English in EFL teaching and learning.

Sadeghi and Heidaryan (2012) investigated whether acquiring discourse markers would influence EFL learners' listening comprehension. The research had 50 males and 50 females of advanced EFL learners whose major was English translation. Fourteen sessions were provided to teach and learn discourse markers with listening comprehension. The research used pretest and posttest to have the quantitative

37

analysis. A control group and an experimental group were involved. The t-test result showed that the experimental group and the control group had a significant difference based on the t-value. The average scores in the experimental group were higher than those in the control group. The research found that after acquiring discourse markers, students started to understand the pragmatic meaning of utterances in addition to the original semantic meanings. They understood that some answers for utterances were not the semantic meaning in a direct way. Thus, discourse markers were important for them to understand utterances and better understand the whole context.

House (2013) investigated whether EFL learners' use of discourse markers improved their pragmatic ability in spoken English. The research tried to find how EFL learners improved their pragmatic competence by using discourse markers, and how EFL learners used discourse markers in conversations. The research studied discourse markers from subjectivity (Benveniste, 1966), intersubjectivity and connectivity point of view. It mainly focused on three specific discourse markers: *yes/yeah*, *so* and *OK*. For *yes/yeah*, in certain situations, EFL learners would use them to uptake messages. Since this discourse marker was so common in daily conversations, learners tended to use them often in all sorts of conditions. For *so*, the result showed that it was used to support elements for speakers in conversations. For *OK*, the research also found that it was the most versatile and broadly applicable discourse markers in spoken English used by EFL learners. It was used easily to achieve different pragmatic functions in dialogues with no bad influence. Speakers tended to like using this discourse marker in all kinds of situations and put it into different positions between utterances. The research concluded that these three discourse markers were commonly used by EFL learners in daily conversations. Their functions changed frequently based on speakers' intentions and topics. Because of the various pragmatic functions of these three markers, EFL learners were able to use

them by selecting their different functions to continue the conversations. However, the research did not compare the use of discourse markers with native English speakers or EFL learners with other backgrounds.

Khatib and Safari (2011) explored discourse markers and their relationship with reading comprehension. The result showed that there was no significant difference between the pretest results and posttest results. Therefore, it concluded that there was no significant relationship between discourse markers acquisition and reading comprehension. However, the research also went deeper to examine each student's test and found that the reason that there was no significant difference was that some answers were not related to the discourse markers directly. Following this condition, researchers interviewed students and examined that students needed discourse markers to fulfill their purpose of expressing things related to various topics. From the answers to the interview, students mainly expressed that discourse markers could help them understand the relationship between the prior utterance and the following utterance. Furthermore, they could help students grasp the main idea of some paragraphs. Thus, it could be seen that discourse markers had partly influenced the answers to questions students needed to reply to.

Yang (2011) pointed out that the importance of discourse markers had paid attention in linguistic research. Therefore, researchers and teachers started to bring discourse markers' acquisition into EFL teaching and learning. Halliday and Hasan (1976) systemic functional grammar was also borrowed into this area to assist the teaching and learning of discourse markers. Pedagogical settings that included teaching discourse markers have been mentioned in several studies (House, 2013). Since each research used different settings to study them, Yang (2011) illustrated that discourse markers in pedagogical settings were still in the infant stage. Teachers and researchers needed to do more analysis and bring theory into practice. Classroom

discourses should be studied further, as discourse markers within different conversation systems were likely different. Thus, future studies should be undertaken for examining the relationship between the discourse markers and language acquisitions.

In conclusion, it can be seen that discourse markers have been studied in SLA within the areas of writing, speaking and reading. Researchers and instructors intended to find out some effective approaches to teach discourse markers for students to realize their importance in communication. Most of the research had a positive result that EFL learners would improve their awareness and use of discourse markers through proper instructions, and discourse markers in SLA would continue to be paid attention to in future studies.

2.4 Studies on spoken discourse markers

The review of the spoken discourse markers is separated into four parts. First, previous research that focused on how native English speakers use the spoken discourse markers (Dehe & Wichmann, 2010; Norrick, 2001; Tagliamonte, 2005) is reviewed. Second, previous research how on non-native English speakers use the spoken discourse markers (Hellermann & Vergun, 2007; Vanda & Peter, 2011) is reviewed. Third, when researchers intended to understand what differences in the use of the spoken discourse markers by non-native English speakers compared to native English speakers. Hence, comparative studies are reviewed (Aijmer, 2004, 2016; Trillo, 2002). Last, studies on discourse markers in spoken data are also relevant to different language styles, such as free indirect style (Blakemore & Gallai, 2014), social informal interactions on the Internet (Prabhumoye et al., 2017), etc. This review also includes these studies.

2.4.1 Studies on spoken discourse markers used by native English speakers

In order to study spoken discourse markers used by native English speakers in spoken English, researchers either studied several spoken discourse markers altogether used by native English speakers or only focused on individual spoken discourse marker. Most studies focused on how native English speakers used certain spoken discourse markers in spoken English; what functions of spoken discourse markers were used by native English speakers and the patterns of how native English speakers used them.

Among these studies, three discourse markers *well*, *like* and *you know* have been studied the most. These studies are reviewed first.

Norrick (2001) explored spoken discourse markers in the oral narrative to try to discover how spoken discourse markers functioned in oral narratives. Since there were a lot of spoken discourse markers, the research chose *well* and *but* to be the objectives of the study. The result illustrated firstly, based on the storytelling materials, *well* could be the spoken discourse marker that aimed at going back to the theme of the story. Secondly, listeners could use *well* as a start to ask questions based on the story. As for *but*, first, the basic function of it should be still a conjunction that the foregoing utterance and the following utterance had something not in common. Thus, by using *but*, the narrator intended to distinguish the former utterance from the upcoming ones. Second, it could also be the spoken discourse marker that the narrator intended to make the theme of the narration back to the main road of the storytelling. Third, it also had the function in the situation where listeners had problems with the story, they would use *but* to make the questions and the whole utterances would be in the story organization. From the analysis of *well*, the researcher summarized that using *well* as a discourse marker was found to be a common phenomenon in storytelling.

Fitzmaurice (2004) investigated three spoken discourse markers: *well*, *you know* and *I mean*. What made Fitzmaurice's (2004) research special was that the researcher took the perspective of analyzing discourse markers from a diachronic point of view. Fitzmaurice intended to figure out how the discourse markers develop themselves. One of the major findings was that these spoken discourse markers definitely had shifted their roles or functions in utterances over time. In history, they were not discourse markers at all. Due to speakers' concern for the actual conduct of interaction, discourse markers were brought into the interaction to oil the wheels of conversational exchange. Selected expressions deployed by speakers to capture what they characterized as the stance or attitude of their interlocutors were used in interactive, as well as subjective and intersubjective functions. In addition, the subjective and intersubjective functions of these apparently fixed phrases were instantiated in both written and spoken registers from the 1650s to the present. Nowadays, their major role as discourse markers has become the interactive function in conversations.

Tagliamonte (2005) explored three common pragmatic markers, *well*, *like* and *but*, used young Canadians in conversation. It aimed at discovering the patterns of young Canadians of using discourse markers in conversations and their functions. The study built a corpus with 26 speakers and 200,000 machine-readable words. Tagliamonte (2005) discovered that the most frequently used spoken discourse marker was *like*, not *well*. The research found that young Canadians usually used *like* in front of a noun phrase, and young people aged 15-16 used this discourse marker the most. However, they could not explain why they used this spoken discourse marker in utterance. They simply used it to express some kind of uncertainty and set examples for previous discourses.

Sakita (2013) specifically investigated the spoken discourse marker *well* carefully as a stance marker used in spoken English. To study *well* deeply, Sakita (2013) had a presentation on different terms of it carefully, including stance, discourse markers and the stance triangle proposed by Du Bois (2007). After carefully investigating its functions in the corpus, two most commonly used contexts were concluded. The first context was the point where the upcoming response was inconsonant with the foregoing discourse and the second context was the topic shift. *Well* negotiated stance and regulated interpersonal relations between conversational co-participants. Therefore, this research finally concluded *well* as a spoken discourse marker with mainly two pragmatic functions. The first one was a stance management in stance divergence, including functions like a response to the prior utterance, marker of insufficiency, preface to a negative or unexpected response, etc. The second one was interpersonal management to regulate interpersonal relations. Sakita (2013) believed that the second function should be the core function of using *well* as a spoken discourse marker.

Waltereit and Detges (2007) used Spanish discourse marker *bien* (*well*) to explore whether this discourse marker in Spanish shared similar functions as in English. The research illustrated two points about discourse markers after studying *well* in both Spanish and English. First, it seemed that discourse markers were more commonly used and appeared across language. They were used more often and could be found everywhere. Second, even though every language may have the same versions of discourse markers, their meanings or functions may be different across languages. Thus, even though the functions of discourse marker *well* in English have been studied, it did not mean that *well* in other languages would have the same functions as in English. Thus, the researchers concluded that both discourse markers and modal particles arise through metonymic semantic change (p. 78). For discourse

43

markers, they linked to contexts where speakers negotiated their further verbal interaction.

So far, it seems that *well* has been studied the most in the previous research. They all have found that native English speakers used *well* as a spoken discourse marker at a high frequency. Different studies adopted different research materials to study *well*, *like* and *you know*, and they all tried to examine the pragmatic functions of each of them.

Besides, some studies focused on other spoken discourse markers. Mostly, each study only investigated one single spoken discourse marker. These studies tried to figure out how native English speakers used them, and to find out what pragmatic functions each spoken discourse marker had in conversation.

Lee-Goldman (2010) explored *no* as a spoken discourse marker in daily conversation. The research was intended to examine how *no* would be the spoken discourse marker in contexts and what its functions were. The conversations used in this research came from two corpora: the ICSI meeting corpus (Janin, 2003, 2004) and the Fisher English Training Corpus (Cieri et al., 2005; Cieri et al., 2004). With Lee-Goldman's study on *no*, the first function of this discourse marker was to get topics or prior utterances serious. Lee-Goldman demonstrated that during the procedure of using *no* to get seriousness on topics or prior utterances, it could also lead to a slight change of the following utterance from the prior utterance. However, the main purpose was to get hearers' serious attention. The second function was to shift topics. There were also other spoken discourse markers that had the same function. The difference was that by saying *no*, speakers had an exact attitude that they wanted to change the topic on purpose. In this function, *no* worked with other lexical items as well. The third function was to manage disagreement. The fourth function was to manage misunderstandings at the logistic level.

Schourup (2011) investigated the use of *now* as a spoken discourse marker with the relevance-theoretic approach. The research was intended to study *now* in a new perspective to find out its functions as a spoken discourse marker. It was found that the first function of *now* was the marker of discourse time. It was close to its semantic meaning. However, in this situation, *now* did not only mean the right moment of the time when the speaker was speaking. *Now* could mean the contemporary time or in recent some time until the moment they spoke. The second function was to be the coherence marker. It did not show semantic meaning, or little semantic meaning, but more in a coherence relation.

What makes Haselow's (2011) research special was that the research explored *then* in the position of the end of the discourse. First, *then* would be a modal particle. Usually, speakers chose a modal verb to express their attitudes or emotions on a thing or an idea. When speakers put *then* in the end, they also tried to express some of their ideas, thoughts, attitudes or emotions with the same function as modal verbs. Second, *then* would be considered as a connection of the previous discourse with its meaning to the following utterances. Thus, through the analysis, the word *then* could be seen as a word that both had its semantic meaning and also had pragmatic meaning in specific conditions.

Dehe and Wichmann (2010) intended to explore two sentence-initial phrases *I think* and *I believe*. The research tried to explore them from the perspective of their prosody. The research used the speaking part of the International Corpus of English and extracted the concordances in relation to these two phrases. Dehe and Wichmann (2010) pointed out that even the research had done the analysis of these two phrases, the real patterns of these two phrases were still undecided. It meant that in the spoken data, people changed their prosody from time to time and there might have been a chance that there would be no patterns or rules. However, with the qualitative

45

examination and extraction of most cases from the data, there still was a basic pattern with a dynamic change in between.

In conclusion, each study reviewed above mainly chose specific spoken discourse markers to examine their functions or features used by native English speakers. With the different situations of experimental studies, information on different frequencies and functions has been discovered.

2.4.2 Studies on spoken discourse markers used by non-native speakers

A few studies only focused on the use of the spoken discourse markers by non-native English speakers to explore how non-native English speakers used discourse markers in spoken English.

First, two studies were found that focused on similar spoken discourse markers. Hellermann and Vergun (2007) explored adult English beginners' spoken English to examine their usage of three discourse markers: *like*, *well* and *you know*. Vanda and Peter (2011) investigated two commonly used spoken discourse markers, *you know* and *I mean*, from the perspective of gender differences. The research tried to examine whether different genders would use these two discourse markers in different ways.

Hellermann and Vergun (2007) pointed out that for English beginners, few spoken discourse markers were used, but 11 out of 17 students used at least one spoken discourse marker. From this point of view, EFL beginners did not know how to use discourse markers. Their chances of using them were quite few. The research suggested that even though teachers might not specifically teach English spoken discourse markers in class, researchers and teachers still needed to pay attention to them and requested students to practice them from time to time. Students needed to target more on spoken discourse markers when they tried to speak English, for spoken discourse markers were significant in conversation.

In Vanda and Peter's (2011) research, for *you know*, the quantitative result showed that there was no significant difference in using it between the two genders. Both genders used it at a high frequency. The research went into deep analysis and found that the functions of this spoken discourse marker used by the two genders were also quite similar. Both genders tended to use it as the functions as follows: to express hesitation; to have a false start; to have personal involvement; to emphasize some topics or points; to have the repetition or to have the explanation of some prior utterances and have the agreement. For *I mean*, the result showed that both males and females also used this discourse marker at a high frequency, in which no significant difference was found between the two genders. Both genders used this spoken discourse marker to express the meanings as follows: to have the topic shift; to elaborate; to explain some prior utterances; to do the clarification; to give some examples or to have contrasts, etc. Thus, there was no difference in using this discourse marker between the two genders.

Tree (2010) reviewed previous research on spoken discourse markers and discovered what had been done and what would be the trend. He pointed out that since recent spoken discourse markers had been studied and confirmed several times that native English speakers used them with their own rules, studies tended to pay attention to the acquisition of the spoken discourse markers in SLA. Research on spoken discourse markers also had results that EFL learners used them when they spoke English, but in different ways or different frequencies compared to native English speakers. The problems of using the spoken discourse markers by EFL learners became a trendy topic that kept being studied since the 21st century. The results in different research would be different based on different participants and various research procedures. In most of the cases, EFL learners were found to use certain spoken discourse markers in different ways. The research pointed out that

47

future studies might focus more on lower-level learners' acquisition of the spoken discourse markers. Tracking EFL learners' acquisition of them in a long term should be considered. More research could focus on how functions of the spoken discourse markers changed during the history of time and how they developed along with the language.

To sum up, several studies have found how non-native English speakers used spoken discourse markers, especially the pragmatic functions of the spoken discourse markers used by non-native English speakers. However, these studies did not have the comparison with native English speakers. More studies appeared to have comparisons with native English speakers to examine the differences in using the spoken discourse markers by both non-native English speakers and native English speakers.

2.4.3 Comparative studies on spoken discourse markers

After the 21st century, the experimental studies mostly focused on comparisons of using discourse markers by native English speakers and non-native English speakers. First, different studies focused on different EFL learners with different backgrounds. Generally speaking, they focused on different nationalities, such as Hong Kong EFL learners (Fung & Carter, 2007; Lam, 2009), Swedish EFL learners (Aijmer, 2011), Belgian native speakers (Buysse, 2012), Turkish EFL learners (Asik & Cephe, 2013), EFL learners from Ireland (Diskin, 2017), etc. Second, most studies mainly focused on how non-native English speakers used spoken discourse markers in terms of their pragmatic functions and differences in the use of spoken discourse markers compared to native English speakers (Aijmer, 2004, 2011; Buysse, 2012; Lam, 2009). They tried to investigate the problems of using the spoken discourse markers by non-native English speakers.

Several studies did not focus on the specific spoken discourse markers. They tried to discover a frequency list of the spoken discourse markers used by non-native

English speakers and compared to native English speaker corpus (Aijmer, 2004; Asik & Cephe, 2013; Trillo, 2002). Their main purpose was to search the top-used spoken discourse markers and how non-native English speakers used them.

Fung and Carter (2007) compared Hong Kong EFL learners with native English speakers to examine whether non-native English speakers and native English speakers used the spoken discourse markers in the same way. The research used the notion of discourse markers as "intra-sentential and supra-sentential linguistic units which fulfill a largely non-propositional and connective function at the level of discourse" (p. 411). Discourse markers play a fundamental role in spoken interaction (Carter & McCarthy, 2006). The study extracted a frequency list and used both quantitative and qualitative analyses. The result found that the use of the spoken discourse markers between Hong Kong learners and native English speakers had a discrepancy in many aspects. Basically, the spoken discourse markers like *and, right, well*, and some other commonly used spoken discourse markers were produced much more by Hong Kong EFL learners. Meanwhile, referential and structural markers were also widely used by EFL learners. The study indicated that a deficiency in using the spoken discourse markers led to a deficiency in pragmatic competence by EFL learners.

Trillo (2002) explored the pragmatic fossilization phenomenon from the perspective of achieving the spoken discourse markers by children and adults. By using the learner corpora, the study aimed at identifying the pattern of evolution of pragmatic markers in non-native English speakers. One of the striking results was that adults used more *look* and *listen* than expected. The research had three conclusions. First, non-native English speakers had a different development pattern in using the pragmatic markers. Second, children usually did not use the pragmatic markers and future research during this stage could be thought of in a less important way. Third, if pragmatic markers would be taught to children in a conscious way, children still had a

chance to achieve the similar use of them as native English speakers. As the findings were shown in this research, Trillo (2002) argued that pragmatic fossilization may influence communication by using few pragmatic markers for both children and adults. A lack of acquisition of the pragmatic markers may lead to long-term pragmatic fossilization for non-native English speakers.

Aijmer (2004) collected the spoken data from EFL learners and had a frequency list of the most commonly used pragmatic markers. The top markers were *I think; sort of; well; I don't know; actually; you know*, etc. (p. 178). The research also listed patterns of the main pragmatic markers used in the data. For example, the patterns related to *well* would be like *well right; well yeah; well I guess; well actually; actually well* (p. 179). Through this study, Aijmer found that EFL learners at a less advanced level might not get used to using these pragmatic markers. Some learners even did not really use or did not really understand their importance of them in daily conversation. Learners used vague and uncertain markers to express uncertainty or hesitation and not for face-saving or to signal politeness. They were also used as strategies when the learners had communication problems.

Asik and Cephe (2013) explored the use of the spoken discourse markers by Turkish EFL learners. The research aimed at discovering Turkish EFL learners' ability to use the spoken discourse markers in spoken English and made a comparison with native English speakers. The result illustrated that the top-used spoken discourse markers were *and*, *umm* and *so*. In most of the cases, EFL learners used the same discourse markers. However, as the result of quantitative data, it showed that learners used fewer spoken discourse markers than native English speakers. There were instances when learners used certain spoken discourse markers more frequently, but only a few could be found, such as *uhh* or *let's*. The research concluded that EFL learners tended to use *uhh* the most. However, in many cases, it could not be counted

as a spoken discourse marker, for its functions were not the same as a spoken discourse marker found in previous research.

Through the study of looking for a frequency list of the spoken discourse markers used by non-native English speakers, it was found that EFL learners in different English-learning backgrounds used different spoken discourse markers in terms of both frequencies and functions. Some spoken discourse markers were used relatively higher than others, such as *like*, *well*, *and*, etc. Thus, several studies followed the early works on discourse markers and tried to investigate some specific discourse markers used by non-native English speakers in detail (Aijmer, 2011, 2016; Buysse, 2012; Diskin, 2017; Polat, 2011). These studies focused on the spoken discourse markers used in a high frequency by native English speakers such as *well* or *like*.

Among these studies, two major findings were discovered. First, some studies found that non-native English speakers used certain spoken discourse markers at much lower frequency in oral communication, or they used fewer functions in terms of each spoken discourse markers under investigation (Aijmer, 2011; Diskin, 2017; Polat, 2011). Second, some studies found that non-native English speakers used certain spoken discourse markers at much higher frequency in oral communication (Buysse, 2012; Lam, 2009).

Aijmer (2011) used Swedish EFL learners to investigate the use of several spoken discourse markers and compared to the use of it by native English speakers. First, the research showed that for both non-native English speakers and native English speakers, the frequency of using *well* was similar, wherein no significant difference was found. Second, the use of *well* in native English speakers had the highest frequency, which was aligned with the result of Biber et al. (1999). The spoken discourse markers *you know* and *I mean* also had high frequencies used by

51

native English speakers, but not as high as the use of *well*. Third, the use of *well* as a spoken discourse marker in terms of its pragmatic functions between non-native English speakers and native English speakers was different. Non-native English speakers treated *well* mainly as two signals. The first was a speech management signal. As this signal, choice-related function and change-related function came in the first place. In contrast, Non-native English speakers tended to use *well* to signal how they wanted this conversation to continue and how they wanted this conversation's topic to be changed to another one. Other functions used as this signal also included marking stages in narratives and quotatives. Second, non-native English speakers treated *well* as an attitude signal. The major roles of this signal were to show speakers' opinions and have a disagreement. Compared to native English speakers, it was found that non-native English speakers used far fewer functions of *well* as a spoken discourse marker compared to native English speakers.

Polat (2011) also focused on the spoken discourse marker *well*. The research explored the use of the spoken discourse markers by immigrant second language learners. The study showed that no use of *well* was found in the learner corpus. The result of this research indicated that there were EFL learners who did not use certain English spoken discourse markers in oral communication at all.

Diskin's (2017) research focused on the use of the spoken discourse marker *like* by non-native English speakers. The data was collected from 42 non-native English speakers who studied English as a second or foreign language in Ireland. Compared to the native English speakers, the frequency of using *like* in non-native speakers was much lower.

In contrast to the studies above, several studies had the results that in terms of the frequencies of the spoken discourse markers, non-native English speakers

overused certain spoken discourse markers or overused certain functions of the spoken discourse markers.

Even though Polat (2011) found that non-native English speakers did not use *well* as a spoken discourse marker at all, non-native English speakers used two of the markers, *you know* and *like*, far more frequently than native English speakers. During the one-year period of the research procedure, the patterns of using different spoken discourse markers were changing a lot. *You know* was the most overused spoken discourse markers in communication, wherein it was used in a high frequency at first but dropped continuously until the end. *Like* was used often with an increase in the whole research procedure.

Buysse (2012) studied *so* as a spoken discourse marker in both EFL learners and native English speakers. The researcher intended to examine the use of *so* by native English speakers, and made a comparison with the use of it by EFL learners. The learner corpus was built with 40 Belgian native speakers. From the quantitative data, it showed that native English speakers used *so* in a lower frequency than EFL learners. Among all the functions of *so* used by non-native English speakers, to indicate a result was the top one, followed by being a prompt and holding the floor. However, it found that native English speakers used fewer *so* in most of the functions except for the function: new sequences. The result also showed that the functions of *so* could be classified based on the relations they indexed in the utterances. Mainly, three types of relations were attested from the original data. The first relation was ideational relation, which was supposed to indicate a result. The second one was interpersonal relation. It could draw a conclusion, prompt or hold the floor. The third relation was textual, which could introduce a summary or a section of the discourse, indicate a shift back to a higher unit of the discourse or as a marker of self-correction, etc. All three functions were found in the learner corpus and were used at high frequencies by EFL

learners. The research concluded that the spoken discourse marker *so* was used in a high frequency by both native English speakers and non-native English speakers. It also had multi-functional uses in different situations. The use of *so* between native English speakers and non-native English speakers was different in both frequencies and functions, wherein non-native English speakers generally overused the spoken discourse marker *so* compared to native English speakers.

Lam (2009) used Hong Kong EFL learners to explore the different uses of the spoken discourse marker *well* between Hong Kong EFL learners and the ones used in the textbooks. The result showed that the use of *well* in learner corpus and in the textbooks was different at a high rate. The first difference was the position of *well*. The EFL learners showed more flexibility in positioning *well* while textbooks only showed in the sentence-initial position. The second difference was the functions of using *well*. Hong Kong EFL learners used *well* in both the textual functions and the interactional functions, which was more than the instances shown in the textbooks. It can be concluded that the textbooks had no specific introduction or guidance on discourse markers and their right use in conversations. The content of discourse markers in textbooks was limited and EFL learners usually did not focus on them. Thus, for using the discourse markers, EFL learners might have a chance to ignore them and when they used discourse markers, they might even not know how to use their functions correctly.

Besides the studies above, other comparative studies mainly explored the different uses or patterns of other spoken discourse markers between native English speakers and non-native English speakers.

Aijmer (2016) tried to discover *anyway* in spoken English by using different corpora to make comparisons. Aijmer used *anyway* as an example to analyze its versatile functions by using three corpora in British English, American English and

54

Philippine English, respectively. It was found that Philippine English used *anyway* the least. Meanwhile, a feature of using it in three corpora was that it co-occurred with other markers, such as *but anyway; oh anyway; so anyway; well anyway; then anyway*, etc. The research also discovered that *anyway* could be a stand-alone marker. It had functions, such as closing a narrative or a topic, abandoning a topic, as a part of a little dialogue to come to an agreement. The research illustrated the fact that spoken discourse markers and their functions were flexible based on different contextual environments.

Dobrovoljc (2017) investigated multi-word discourse markers (MWDMs) by using corpus-driven semi-automatic approach and qualitative approach to identify the most frequently used multi-word discourse markers in spoken Slovene. The result had little difference from the previous research with limited selection. In the 173 items, 120 items were a combination of 2 words. 95 items were a combination of function and content words. In terms of the semantic features, the most used MWDMs were vagueness, dialogue and formation.

In conclusion, what has been discovered in all of the comparisons reviewed above was that there were differences in using the spoken discourse markers between native English speakers and non-native English speakers. The frequency of the spoken discourse markers investigated in most of the research was different between native English speakers and non-native English speakers. The pragmatic functions of certain spoken discourse markers used by non-native English speakers were also different from native English speakers. In most studies, non-native English speakers did not have a comprehensive acquisition of the functions of the spoken discourse markers under investigation. Because of all these experimental studies, it was known by the researchers that these differences influenced non-native English speakers' oral communication and their pragmatic competence. Although non-native English

speakers used some spoken discourse markers in oral communication, it was found that some spoken discourse markers were overused, while some spoken discourse markers were underused compared to native English speakers. In the meantime, it was found that the pragmatic functions of some spoken discourse markers used by non-native English speakers were different from the ones used by native English speakers in communication. Moreover, these various pragmatic functions made each spoken discourse marker essential to the communication because the functions influence the pragmatic understanding between the speaker and the hearer. It was thus conceivable to see, as cited by the previous research (Aijmer, 2011, 2016; Fung & Carter, 2007), that an inappropriate use of spoken discourse markers may lead to a deficiency in pragmatic competence in communication.

2.4.4 Studies on discourse markers in special types of spoken data

The experimental studies that have been reviewed above achieved their results in the use of the spoken data in the genre of daily conversations (Aijmer, 2011; Diskin, 2017). Some other studies were found to examine the spoken discourse markers in other types of social interactions, such as free indirect style and communication online (Blakemore & Gallai, 2014; Prabhumoye et al., 2017).

Blakemore and Gallai (2014) investigated spoken discourse markers in free indirect style and in their interpretation. The study aimed at discovering how free indirect style and interpreting contributed to communication, and how spoken discourse markers in both styles influenced communication. In the free indirect style, the spoken discourse markers, such as *well*, were found to express the procedural meaning. It was aligned with its use in other types of texts. It was as a means of reducing the possibility that hearers would continue to interpret the prior utterances again. However, sometimes, this function used by speakers would not be discovered totally by hearers. Thus, hearers may or may not find the same hints or signals. The

intention for speakers to use the spoken discourse markers, such as *well* or *anyway*, was to change the environment of the whole utterances in the specific situation. In interpreting, the spoken discourse markers, such as *so* and *well*, tended to be used more often. Their roles in utterances were challenging for interpreting, as interpreters should be loyal to the original script. During the procedure of interpreting, interpreters would change the spoken discourse markers in certain instances, or they would find another spoken discourse marker that was supposed to be in the same sense between two languages.

Taboada and Gomez-Gonzalez (2012) examined concession discourse markers across both speaking and writing in English and Spanish. The research was intended to explore the coherence relationship and how discourse markers signaled it. It was found that the discourse markers that signaled concession relations in both written and spoken forms were similar but in different frequencies. The most commonly used discourse markers were *but; although; while; however*, etc. The frequency of using the same discourse marker in speaking was much lower than the one in writing. Compared both languages, the use of the Spanish written discourse markers was higher than the English written discourse markers. In contrast, the use of the English spoken discourse markers was higher than the use of the Spanish spoken discourse markers. Meanwhile, the functions of using the same discourse marker in both genres were sometimes also slightly different. In the written form, the functions of the discourse markers would qualify writers' opinion, or dismiss potential disagreement from writers' point of view. In spoken form, they were used to indicate a contrast in different situations, or speakers intended to correct certain potential misunderstandings in the utterances.

Tree (2015) was interested in the informal language online and explored the spoken discourse markers used in the informal language online. The research was

intended to find out whether the spoken discourse markers that were unique in communication were also massively used in informal typing, how they were used and their features in informal typing. It was found that *well* was used the most to express speakers' attitudes towards different aspects, like speakers did not agree with the prior discourse. Other spoken discourse markers were also used, but not so much as *well*. Tree pointed out that on the Internet, the linguistic markers in the informal interaction had more forms or types than in any other settings in the real world. People could only use a symbol or a single phonological sound to represent their meanings other than semantic meanings. With the same interest, people communicated with their short-cut unique language and these linguistic markers could only appear in this setting. Moreover, since the conversations happened every day, more interesting linguistic markers could appear and add to the new communication.

A brief conclusion is presented to summarize this part of the review. The results of studying the English spoken discourse markers used by native English speakers have pointed out three ideas. First, the spoken discourse markers used by native English speakers were different from the written discourse markers, wherein the spoken discourse markers were generally not used in English writing. Second, native English speakers used discourse markers at a high frequency in oral communication. Third, various pragmatic functions of the spoken discourse markers were used in conversational interactions. Unlike native English speakers, non-native English speakers used the spoken discourse markers differently in both frequency and pragmatic function. Several studies have shown results that non-native English speakers overused some discourse markers, while other studies illustrated that non-native English speakers underused the spoken discourse markers. Furthermore, most studies concluded that non-native English speakers mastered a limited amount of

spoken discourse markers in their limited pragmatic functions, which led to a deficiency in pragmatic competence in oral communication.

As stated above, studies on spoken discourse markers have been conducted all around the world. In recent years, experimental studies have focused on the comparisons in using the spoken discourse markers between native English speakers and non-native English speakers. Studies had different perspectives to do the research, i.e., discourse markers in conversation or in speech; discourse markers with single particles or multi-word items; the same discourse marker in different languages, etc. Even though there was still a controversy in the basic issues of the discourse markers, comparative studies on differences in the use of the discourse markers between native English speakers and non-native English speakers have had benefits for future studies.

After the reviews of the studies on discourse markers in the world, studies on both Thai discourse markers and English discourse markers in Thailand will be presented in the next part.

2.5 Studies on discourse markers used by Thai EFL learners

In this part, the studies on discourse markers by Thai EFL learners are reviewed. The use of Thai discourse markers by native Thai speakers will be reviewed first, followed by the review of the use of English discourse markers by Thai EFL learners.

2.5.1 Studies on Thai discourse markers used by native Thai speakers

The previous studies on the use of Thai discourse markers by native Thai speakers are reviewed first. The review of these studies is to examine how native Thai speakers used Thai discourse markers in Thai contexts. Moreover, the review may provide a perspective on how L1 Thai influenced the use of English discourse markers by native Thai speakers.

Several studies focused on Thai discourse markers used by native Thai speakers (Hiranras, 2011; Kittopakrankit, 2018; Simma, 2014; Wutthichamnong, 2016). These studies examined the functions of individual Thai discourse marker used by native Thai speakers.

Hiranras (2011) examined the Thai discourse marker ʔaw (เอา). It mainly discussed how this word became a discourse marker from its original part-of-speech as a verb from a grammaticalization point of view. It was found that as a discourse marker, ʔaw (เอา) mainly had two functions: adding a concept into a sentence and giving a turn to the speaker of the conversion. The study discovered that ʔaw (เอา) changed its verb part-of-speech into a discourse marker through four stages from a grammaticalization point of view. Cognitive processes, including metaphor and metonymy, contributed to the grammaticalization path for the change. It was concluded by the study that ʔaw (เอา) was used as a verb and a discourse marker in Thai discourse.

Simma (2014) examined the use of sûŋ (ซึ่ง) in Thai discourse. It was found that in contrast to the previous findings where sûŋ (ซึ่ง) was used as an additive and casual meanings, it was used as a topic marker the most, in which sûŋ (ซึ่ง) was a component or all information expressed in the previous utterances. Simma found that sûŋ (ซึ่ง) was used in Thai discourse with no or little syntactic meaning, but to connect utterances. It was thus believed that sûŋ (ซึ่ง) was used as a discourse marker in Thai discourse.

Wutthichamnong (2016) examined the English loan word *Okay* (โอเค). It aimed at analyzing the pragmatic function of the Thai word *Okay* (โอเค) as a discourse marker in Thai conversation and comparing to the functions of the English word *Okay* (โอเค)

as a discourse marker. It was found that eleven functions of the Thai discourse marker *Okay* (โอเค) were used in Thai discourse. Compared to the use of the English discourse marker *Okay*, the differences were that the Thai discourse marker *Okay* (โอเค) was used for confirming and closing, where no case of using these two functions was found in the English discourse marker *Okay*. It indicated that the Thai discourse marker *Okay* (โอเค) was used often in various pragmatic functions because it was an English loan word that influenced native Thai speakers to use it in Thai discourse.

Kittopakrankit (2018) compared the discourse markers between Thai and Chinese languages to examine the features of Thai discourse markers. It was found that Thai discourse markers were mainly from modal particles, e.g. khrab (ครับ) and kha(ค่ะ). Instead of the interpersonal purposes of using the discourse markers in other languages, the use of Thai discourse markers by native Thai speakers was mainly for the purpose of politeness and intimacy.

Based on the findings above, it can be seen that Thai discourse markers were not the same as English discourse markers. It may concern with the fact, as illustrated in previous research (Kittopakrankit, 2018), that the Thai culture and the interactional purposes of native Thai speakers differed from the ones in other languages.

2.5.2 Studies on English discourse markers used by Thai EFL learners

Some studies focused on written English discourse markers used by Thai EFL learners (Chotiros, 1999; Jangarun & Luksaneeyanawin, 2016; Prommas & Sinwongsuwat, 2011; Sitthirak, 2013).

Chotiros (1999) compared the use of English contrastive discourse markers (ECDMs) and Thai contrastive discourse markers (TCDMs) by Thai EFL learners. The study revealed that Thai EFL learners used the same cognitive pattern to express the pragmatic meaning by using CDMs in both languages. However, it was found that

there was no a one-to-one correspondence between the CDMs of the two languages, resulting in a different usage of CDMs in both languages.

Sitthirak (2013) also investigated contrastive discourse markers (CDMs) used by Thai university students at different levels and made a comparison to native English speakers. The result firstly showed that Thai EFL beginners and Thai intermediate-level EFL learners were able to identify CDMs. When they chose CDMs to use, they usually examined the semantic meaning of the utterances and then examined the coherence. For native English speakers, when they were asked to choose more than one CDMs in a single utterance, they had struggles, which was a normal phenomenon according to the research. The research concluded that for CDMs used in both speaking and writing, EFL learners chose CDMs according to the specific situation. Basically, they chose the informal word *but* in speaking and formal word *nevertheless* in writing.

Prommas and Sinwongsuwat (2011) investigated the use of discourse connectors in argumentative essays and made comparisons between Thai EFL learners and native English speakers. Eight types of discourse connectors were examined in this research. The frequency result showed that Thai EFL learners used discourse connectors more frequently than native English speakers. However, native English speakers used more types of discourse connectors than Thai EFL learners. In the top list of used discourse connectors, 4 of them were the same but in different orders used by native English speakers and Thai EFL learners. When the researchers examined the use of each discourse connector deeply, they found that native English speakers used more functions of each discourse connector. Moreover, Thai EFL learners might use some discourse connectors incorrectly, or they mistakenly used some discourse connectors to substitute another one.

Jangarun and Luksaneeyanawin (2016) had a comparative study on the discourse connectors in argumentative essays written by three groups of students: American undergraduate students, Thai with high-English exposure and Thai with low-English exposure. It examined the similarities and differences among the three groups of students. This research categorized the written discourse connectors into five genres: Additive, Adversative, Casual, Temporal and Continuatives. The result discovered that the use of two of the categories, which were Casual and Temporal, had a significant difference between Thai students and American students. The research pointed out that the different uses of certain discourse connectors may be the effect of interlanguage processes by Thai students. In the meantime, Additive has been used the most among all the three groups. It was found that when students used Additive *and*, there were pragmatic functions other than its original additive meaning in sentences. This research reflected that, to some extent, Thai students used written discourse connectors similar to American students, while there were differences in using the written discourse connectors in some categories.

Besides the studies of written discourse connectors, few studies focused on the spoken discourse markers used by Thai EFL learners (Arya, 2020; Nookam, 2010; Tantiwich & Sinwongsuwat, 2019).

Arya (2020) examined the distributions of the spoken discourse markers that were used most frequently by Thai EFL learners. The result illustrated that the most-used spoken discourse markers by Thai EFL learners were *and*, *OK*, *but* and *so*. As a further investigation on the functions of the spoken discourse markers used by Thai EFL learners, it was found that the referential purposes, which mainly connected the relationship of two utterances, were used the most. It seemed that Thai EFL learners lacked the use of spoken discourse markers for the interpersonal and cognitive purposes. Arya (2020) pointed out that it was clear to see that there was a deficiency

63

in using the spoken discourse markers by Thai EFL learners. In addition, as interpersonal and cognitive purposes of the spoken discourse markers affected the pragmatic understanding in oral communication, the research indicated that a pedagogical urgency should be thought of to improve the awareness of using the spoken discourse markers for Thai EFL learners.

The same as Arya (2020), Nookam (2010) also examined a frequency list of the top-used spoken discourse markers by Thai EFL learners. It was found that the top-used spoken discourse markers were *and*, *OK*, *but* and *so*, which were exactly the same as the findings in Arya (2020). Meanwhile, the study pointed out that the interactional functions were rarely used by Thai EFL learners.

Tantiwich and Sinwongsuwat (2019) explored the use of *yes/no* tokens in spoken interaction among Thai university students. It followed the previous research that examined *no* as a discourse marker (Lee-Goldman, 2011) to investigate both *yes* and *no* produced by Thai EFL learners from both discourse markers and non-discourse markers' perspectives. Based on the data collected in the research, six types of functions of *yes* were identified used by Thai university students. The first two functions, to indicate acceptance and to have the confirmative response, were used much higher than the other four functions. The other four functions were to have positive alignment, to make the acknowledgment, to have the topic shift, and to have the self-confirmation. For the result of *no*, three major functions were found. The first function was to express speakers' disconfirmation or disagreement. The other two functions were to have the restatement and to have a negative alignment. Based on similar studies abroad (Lee-Goldman, 2010; Schegloff, 1992, 2001), the research had a comparison of the use of *yes* and *no* produced by both Thai university EFL learners and native English speakers. The result illustrated that, for *yes* token, there were functions that were not used by Thai university students, such as to show agreement

and to signal incipient speakership. For *no* token, more functions were found in other studies but were not used by Thai university students. It indicated that Thai EFL learners had a deficiency in using both *yes* and *no* tokens as the discourse markers.

Based on the reviews of the studies on the English discourse markers used by Thai EFL learners, it can be seen that Thai EFL learners, to some extent, had difficulty in using the English discourse markers in both writing and speaking. As pointed out above, few studies on English spoken discourse markers used by Thai EFL learners were found.

In the next part, since this research includes six English spoken discourse markers, their pragmatic functions in the previous research will be reviewed one after another in detail.

2.6 Pragmatic functions of spoken discourse markers in previous research

This section is separated into three parts. The first part is to have a brief introduction of the methodology of identifying the functions of discourse markers in the previous research. The second part is to have an introduction to Brinton's framework of functions of discourse markers (1996, 2008). The third part is to list major pragmatic functions that have been found in the spoken form used by native English speakers (Anderson, 2000, 2001; Diskin, 2017; Müller, 2005; Polat, 2011; Popescu-Belis & Zufferey, 2011; Tagliamonte, 2005).

2.6.1 Methodology of identifying the functions of discourse markers in the previous research

Different previous research used different methods to analyze the functions of discourse markers. Based on the review of the previous research, two main methods have been identified.

The first method is that the researchers analyzed the functions of discourse markers based on a theoretical framework (Aijmer, 2011, 2016; Blakemore & Gallai, 2014; Der, 2010; Fitzmaurize, 2004; Khedi et al., 2013; Sakita, 2013; Schourup, 2011). It was found that Relevance Theory was used the most to be the framework for function analysis in early work. (Anderson, 2001; Blakemore, 1987, 1992; Blakemore & Gallai, 2014; Blass, 1990; Jucker, 1993; Schourup, 2001, 2011; Watts, 1988).

With the development of the studies on discourse markers, some researchers built new frameworks for analyzing the functions of discourse markers (Brinton, 2008; Fung & Carter, 2007; Schiffrin, 1987). These frameworks were specifically used to discover the functions of discourse markers in the data of non-native English speakers, such as Brinton's framework (Aijmer, 2011; 2016; Der, 2010), Schiffrin's framework (Redeker, 1990), etc.

Schiffrin's framework (1987) has been briefly illustrated in 2.1. Brinton's framework (2008), which is used as the framework for identifying the functions of spoken discourse markers in this research, will be illustrated in the next part. Therefore, this part illustrates the Relevance Theory first.

Relevance Theory was firstly proposed by Sperber and Wilson in 1986. It has been used in the interpretation of language to help researchers analyze and understand utterances. The theory adopted ideas from cognitive linguistics. Sperber and Wilson (1986) believed that in oral communication, communicators might convey more information than the literal meaning of the utterances. From the surface of the verbal communication, the speaker intended to make the hearer understand the information beyond the literal meaning of the utterances. Hence, the literal meaning and the information beyond the literal meaning were somehow "relevant" to each other. So, from the speaker's point of view, two intentions were discovered. First, the speaker intended to inform something to the hearer (the informative intention). Second, the

speaker intended to inform his/her informative intention to the hearer (the communicative intention). The hearer interpreted the literal meaning with the "relevant" information to discover the real intention of the speaker. The hearer may have different conclusions based on the literal meaning of the utterances spoken by the speaker. Thus, the theory proposed two principles to follow. First, human cognition tends to be geared to the maximization of relevance. Second, every utterance conveys the information that is relevant enough for it to be worth the hearer's effort to process it and it is the most relevant one compatible with the speaker's ability and preference. The two principles suggest that the hearer tends to choose the most relevant interpretation of the utterances from the speaker cognitively with an effort that the hearer believes to be worth interpreting. It is overtly to see that during the interpretation, both the speaker and the hearer are involved in the procedure of cognition and interaction.

Some researchers believed that the Relevance Theory was able to interpret the intention of speakers and the interaction between the speaker and the hearer (Blakemore, 1992). Jucker (1993) argued that it was the only theory that fit to analyze the discourse marker *well*. Many researchers followed Relevance Theory to study the functions of discourse markers used by native English speakers from the 1980s to the early 2000s (Andersen, 2001; Blakemore, 1987, 1992; Blass, 1990; Jucker, 1993; Watts, 1988). In their research, after illustrating the theory, the researchers usually analyzed each function of certain discourse markers one after another in the given context based on the central idea of the theory. Although there has been criticism on the issue of using this theory to analyze the pragmatic use of language (Borderia, 2008) and some research in recent years adopted other frameworks to study discourse markers from new perspectives, Relevance theory still has contributed to finding out the basic functions of discourse markers in a long time.

Besides Relevance Theory, other frameworks have also been used. Lee-Goldman (2011) used principles of conversation in conversation analysis to study the functions of *no* as a discourse marker. Sakita (2013) used Du Bois's dialogic framework (2007) to study how the discourse marker *well* expressed stance in the conversation. Aijmer (2016) used a constructional approach to study *anyway* as a discourse marker. It can be discovered that the researchers used different frameworks to investigate an individual discourse marker's functions from different perspectives in recent years.

The second method is that the researchers reviewed the findings on the functions of the individual discourse marker illustrated in the previous research and used the findings of the functions of this individual discourse marker to analyze the use of it by EFL learners in a combination of the corpus-driven approach (Aijmer, 2011; Alghamdi, 2014; Asik & Cephe, 2013; Buysse, 2012; Diskin, 2017; Vanda & Peter, 2011). Diskin (2017) named it a literature-driven approach, while others named it a top-down approach (Buysse, 2012; Müller, 2005). This method was mainly used when functions of certain discourse markers under investigation had been determined in the previous research with no or little disagreement. The researcher firstly reviewed the major findings of the previous research on the same discourse marker(s) that had been attested in the data of native English speakers, and then had a conclusion of the functions that the previous research had discovered and acknowledged by other researchers. The researchers sorted out the previous findings, categorized them and listed all the functions that had appeared in the previous studies. Based on the conclusion of the functions in the previous research, the researchers used the corpus-driven approach to analyze their own data. Usually, this method was used in research that investigated an individual discourse marker in the data of non-native English speakers (Alghamdi, 2014; Asik & Cephe, 2013; Buysse, 2012; Diskin, 2017).

There are some other methods found in the previous research. In Polat's research (2011), the researchers had two raters to identify the functions of *you know* and *like* as discourse markers. Müller (2005) only used the corpus-driven approach (bottom-up approach) to identify the functions of discourse markers, although he also claimed that the previous findings would be a reference. Müller (2005) argued that from the corpus-driven analysis, evidence from the data took precedence over theoretical constructions.

In conclusion, in the late 20th century, researchers used certain theories that related to pragmatic linguistics to analyze the functions of discourse markers. Most of the data were from native English speakers. The major findings of these studies provided foundations for the research on non-native English speakers' use of discourse markers since the 21st century. When researchers attempted to study the discourse markers used by non-native English speakers, they used the literature-driven approach (the top-down approach) combined with the corpus-driven approach. Researchers pointed out that multiple approaches had been adopted in the study of discourse markers (Aijmer, 2011; Bolden, 2015; Der, 2010).

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

Brinton used to study discourse markers from the diachronic perspective to investigate the change and the development of discourse markers in English, Latin and some other languages (1996). In 2008, Brinton presented a more detailed framework that had a whole scenario of the functions of discourse markers that would appear in communication. The framework follows Schiffrin's five planes of talk (1987) by adopting Halliday's systemic functional linguistics model of language (1979) and Relevance Theory (Brinton, 2008; Aijmer, 2011). Besides, Brinton also consulted the previous findings of empirical studies to construct the framework (Bazzanella, 2006; Brinton, 1996; Erman, 2001). The framework indicates Brinton's

69

belief that discourse markers have little or no propositional meaning but served a pragmatic or procedural purpose. There are two main domains to separate the functions of discourse markers. The first domain is from the interpersonal perspective. Brinton (2008) argued that this domain of functions was an intrinsic feature of discourse markers. It is believed that discourse markers are used as vehicles to show the relationship between the speaker and the hearer. This idea is aligned with the exchange structure and participation framework that were proposed in Schiffrin's framework (1987). There are many kinds of relationships in which the speaker and the hearer involve in communication. Brinton (2008) briefly separated the interpersonal functions into two groups: the subjective functions and the interactive functions. For one thing, the speaker tends to express their own attitudes towards people, things or topics subjectively, such as showing the speaker's personal opinions, showing the speaker's stance, holding the floor from the speaker's point of view, etc. For another, the speaker needs to interact with the hearer by inserting certain discourse markers in utterances, such as using *well* as a hedge, using *I mean* to show the emphasis of the following utterances, etc. Each discourse marker has different functions to intrigue the relationship between the speaker and the hearer. The same discourse marker can also be used to intrigue different relationships with different functions. The second domain is from the textual perspective. It reveals the relationship between the prior utterances and the following utterances. Brinton argued that textual functions of discourse markers were able to mark various kinds of boundaries of discourses in the communication, such as topic shifting, a new topic opening, etc. Textual functions also signal turn-taking in oral communication. This domain is aligned with the action structure and information state in Schiffrin's framework (1987). Textual functions are also dependent on the relevance of specific

discourse markers and contexts. Brinton's framework of functions of discourse markers is presented in Figure 1.

Figure 1: Brinton's framework of functions of discourse markers (2008)

```
            ┌─────────────────────────────────┐
            │  functions of discourse markers │
            └─────────────────────────────────┘
                    │                   │
        ┌───────────────────────┐  ┌──────────────────┐
        │ interpersonal functions│  │ textual functions│
        └───────────────────────┘  └──────────────────┘
```
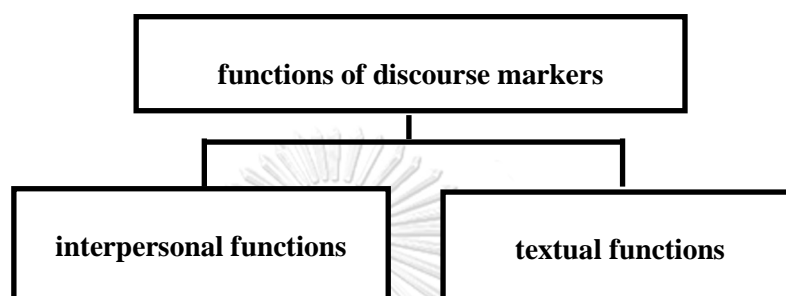
The functions under interpersonal functions and textual functions are shown in Figure 2 and Figure 3, respectively.

Figure 2: Interpersonal functions in detail adapted from Brinton (2008)

```
                  ┌────────────────────────┐
                  │ interpersonal functions│
                  └────────────────────────┘
                   │                      │
          ┌────────────┐           ┌────────────┐
          │ subjective │           │ interactive│
          └────────────┘           └────────────┘
```

stance	opinions	hold the floor	hedge	emphasis	shared knowledge	implicature
e.g.	e.g.	e.g.	e.g.	e.g.	e.g.	e.g.
well	*so*	*so*	*well*	*I mean*	*you know*	*well*

71
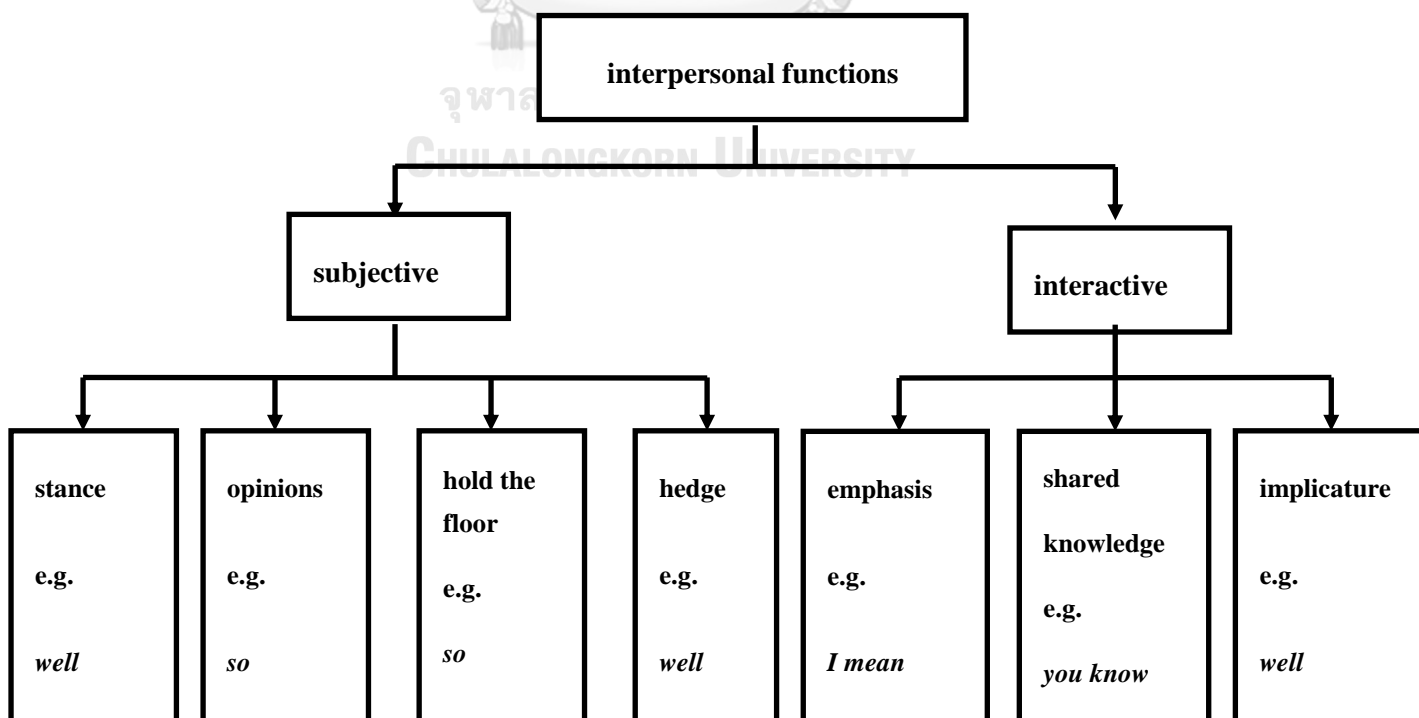
Figure 3: Textual functions in detail adapted from Brinton (2008)

textual functions

| topic shift e.g. *well* | a repair e.g. *like, I mean* | a quote e.g. *like, you know* | a request/ question e.g. *so* | more explanation e.g. *like* | a result e.g. *so* | a gap filler e.g. *you know* |
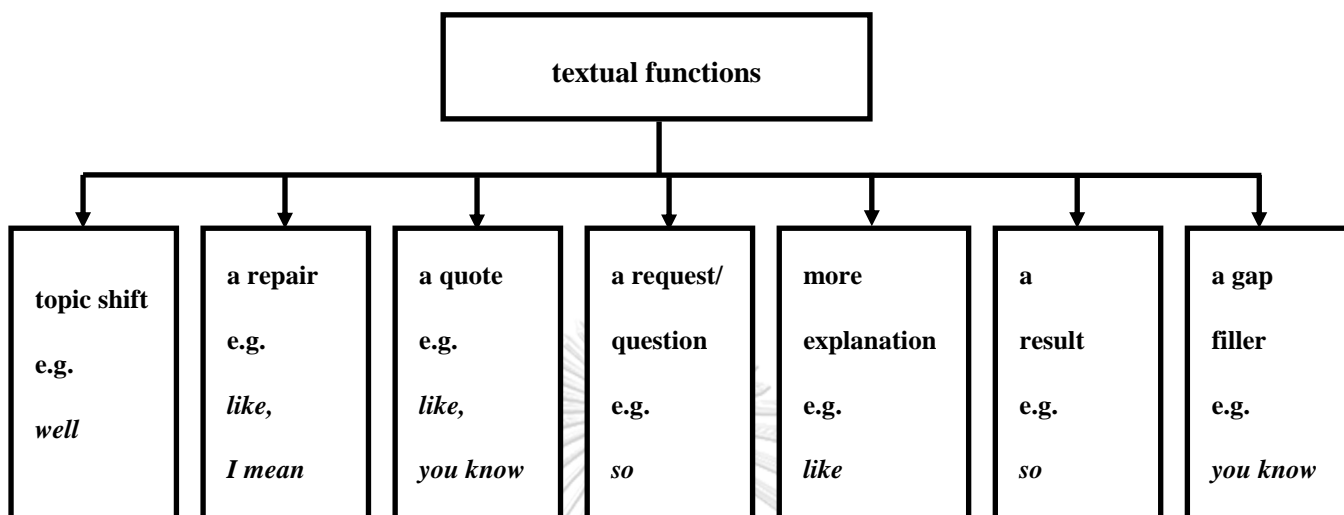
Based on this framework, Brinton used *I mean* (2008, 2010) to further explain the specific functions under each category. Through the analysis of the data and review of the previous studies, the functions of *I mean* as a discourse marker were categorized into expressing subjective stance under interpersonal functions and two textual functions. In the subjective stance functions, there are four specific stances: showing a misinterpretation of the speaker; showing an explanation of the speaker's points subjectively; showing the speaker's attention to something subjectively and showing the speaker's hesitation in the on-going interaction. In the textual functions, there are two specific functions: an explanation of the prior utterance(s) objectively and marking a topic change.

Brinton (1996) argued that it was possible for a discourse marker to have more than one function simultaneously, developed from Ostman (1982) and Schiffrin (1987). However, it was believed that discourse markers generally should have a focus on a concrete context (Brinton, 1996; Ostman, 1982), wherein the interactive purpose may

be considered central. It can be interpreted that there should be a primary function of a discourse marker in a specific context, especially if the interactive purpose can be found under the circumstances. In the recent experimental studies, the researchers mainly used frameworks proposed by previous researchers or corpus-driven approach to identify the main function of a discourse marker in different contexts (Brinton, 2010; Buysse, 2012; Diskin, 2017; Fung & Carter, 2007). Hence, this research tries to analyze the main function of each spoken discourse marker in each specific context performed by Thai EFL learners by the guidance of Brinton's (2008) framework with an assistance of the Relevance Theory.

To sum up, Brinton's (2008) framework is not only constructed on the basis of previous theories and research, but also guides researchers to analyze the functions of discourse markers in future studies (Aijmer, 2011, 2016; Der, 2010).

2.6.3 Functions of *like*

To examine the specific functions of each discourse marker in this research, the functions of each discourse marker that have been attested in previous research on native English speakers are reviewed and presented in this part. Each spoken discourse marker's pragmatic functions can be separated into textual functions and interpersonal functions according to Brinton's (2008) framework. Moreover, each spoken discourse marker's unique functions are concluded based on the findings of the previous research.

It was found that the spoken discourse marker *like* had seven textual functions and one interpersonal function based on the findings of the previous research (Diskin, 2017; Müller, 2005; Popescu-Belis & Zufferey, 2011; Tagliamonte, 2005), illustrated in Table 2.

Table 2: Pragmatic functions of the spoken discourse marker like based on findings of previous research

Like		Searching for the following words
As a	**Textual functions**	To exemplify
discourse		As a repair
marker		To have a quotation
		More information/explanation
		As an approximator
		Marking a focus on new information
	Interpersonal function	As a mitigator

First, it is used to search for the following words. When this function is used, there is usually a long pause after *like*, indicating that the speaker is in the procedure of searching the right words. Sometimes, the following utterance is incomplete or unclear due to the failure of searching. For example, he started *like…*kind of doing…

Second, it is used to exemplify certain prior information. It is used to present detailed information or example of a prior utterance. The example and the prior utterance should not be the direct superordinate and subordinate relationship. For example, they had to act extremely *like* starting to vomit.

Third, it is used to be a repair. It is used when the speaker corrects himself. For example, it's just funny how he got his…*like* how he sometimes leaned there…

Fourth, it leads to a quotation. For example, he was *like*…oh, I'm the captain. Listen to me.

Fifth, it leads to provide more information or additional explanation. The more information or the additional explanation is to supplement the prior utterance. They are not new information. For example, he's medium…*like* not so good.

74

Sixth, it is used as an approximator. It is used as this function to show that the following utterance may not be totally correct or it is a loose fit. The following utterance is usually a single word or short phrases. For example, he was *like* a movie director or something.

Seventh, it marks a focus on new information. It is used when the speaker intends to make the hearer know that the following utterance is new information, and it should be important to be paid attention to. It is worth mentioning that the information provided after *like* should be totally new, which never appears before *like*. For example, I think we met in *like* high school.

The one interpersonal function is listed as follows.

It is used as a hedge or a mitigator to mitigate a potential argument in communication. It can be either sentence-initial or sentence-final. For example, do you think you could do it a little bit faster, *like* you know I've got no money here…

2.6.4 Functions of *so*

It was found that the discourse marker *so* had seven textual functions and five interpersonal functions based on the previous research (Blakemore, 2014; Buysse, 2012; Tagliamonte, 2005), as illustrated in Table 3.

Table 3: Pragmatic functions of the spoken discourse marker so based on findings of previous research

so **As a discourse marker**	**Textual functions**	Marking a result/consequence
		Topic shift back
		Leading a summary
		A transition sequentially
		Marking a start of a narration
		Self-correction
		More explanation

		Marking a question/request
		Marking opinion/feeling/stance
Interpersonal functions		Prompting the hearer to speak/an implied result
		Marking a start of a new topic for involvement
		hold the floor

The seven textual functions are listed as follows.

First, it marks a result or a consequence. When this function is used, the prior utterance should not be the direct cause of the following utterance. The following utterance is a result, or a consequence of a deduction based on the prior utterance. There can be several prior utterances to lead to the result or the consequence. For example, he won't come, *so* let's start.

Second, it leads to a topic shift back situation. For example, yeah, I like it…it was funny and there is the actor I like a lot…and you see the scene with…*so* I enjoy the movie very much.

Third, it is used for the transition sequentially. For example, another guy walked in, *so* he sat down and Charlie saw him….*so* he wanted to talk to him.

Fourth, it is used to mark a self-repair. For example, she wants to have her brain washed…*so*, she has a brainwash.

Fifth, it leads a summary of several prior utterances. It is worth mentioning that to distinguish this function from the first function, namely marking a result or consequence, it is essential to examine whether the utterance preceded by *so* is a conclusion of a series of prior utterances or a result that can be deducted from the prior utterances For example, he just grabbed some food, even did not talk to me, and just gone…*so* he just left. The function of the spoken discourse marker *so* in this example is considered as a summary of the prior utterances is because the utterance

he just left preceded by *so* concludes all the actions in the prior utterances, including *grabbed some food*; *did not talk to me* and *just gone*. Besides, the action *he just left* has been mentioned in the prior utterances by the words *just gone*. It is thereby inappropriate to make the utterance *he just left* as a result of the prior utterances.

Sixth, it leads more explanation or elaboration. For example, when I came here I didn't plan to study Italian…*so*, I was gonna do linguistics but cos you've got to choose three subjects and what not…

Seventh, it marks a start of a narration. There should be a new narration following it. For example, *so* what happened when you went out was that they came into the house, and grabbed a bunch of beer…

The five interpersonal functions are listed as follows.

First, it leads a question or a request from the speaker. For example, *so* what do you think of this film…

Second, it marks speaker's opinion, feeling or stance towards something. For example, *so* that's weird.

Third, it is used to prompt the hearer to speak, or to provide an implied result. It is usually inserted at the end of the utterance. For example, I didn't see it that way, *so…*

Fourth, it is used to hold the floor. For example, *so…*the management types of jobs should be presented by…

Fifth, it marks a new topic. For example, *so* I heard Alice got fired…

2.6.5 Functions of *well*

Well is one of the discourse markers that has been studied the most. It was found that it had five textual functions and seven interpersonal functions based on the previous research (Blakemore, 2014; Brinton, 1996; Fung & Carter, 2007; Lam, 2009; Müller, 2005; Norrick, 2001; Sakita, 2013), as illustrated in Table 4.

Table 4: Pragmatic functions of the spoken discourse marker well based on findings of previous research

well **As a discourse marker**	**Textual functions**	Searching for the right words
		Self-correction
		Marking a new topic
		More explanation/information
		Leading to a conclusion
	Interpersonal functions	Indicating an incomplete answer
		Indicating the unexpected answer
		A face-threat mitigator
		Marking a request (a new start)
		A qualifier to mark agreement/reinforcement
		Marking a dispreferred response
		Leading or expressing personal stance/feeling/opinion

The five textual functions are listed as follows.

First, it is used to search for the right words. For example, then he ends on the street of…*well*…I guess New York.

Second, it is used for self-correction. For example, they say that…*well*…he says that…

Third, it marks a new topic. For example, we just ate lunch…*well*…I feel tired. Let's go home.

Fourth, it leads more explanation or information. For example, I talked to my supervisor…*well*…he's a lecturer.

Fifth, it leads to a conclusion. It is used when the speaker wants to end the conversation. For example, *well*…that's it I guess.

The seven interpersonal functions are listed as follows.

First, it is used to be an implicature of insufficiency of an utterance. It is usually used when the speaker answers to a prior question. For example, …have you done it yet… *...well*…yes, maybe.

Second, it marks an implicature of an unexpected answer. It is used when the speaker indicates that the following answer to the prior question may be not so expected as the hearer's intention. For example, *well...*let's just say he's not so good.

Third, it is used as a face-threat mitigator. It is used at this point to avoid a potential argument. For example, A: what is wrong with you…B: *well...*I just told him a little…really.

Fourth, it is used as a qualifier to mark an agreement or reinforcement. For example, *well*…thank you for your understanding.

Fifth, it marks a dispreferred response. It is used when the speaker replies to the previous speaker with disagreement. It is therefore essential to examine whether there is a disagreement preceded by *well* against the utterance by the prior speaker. For example, *well...*actually he's not so right about it.

Sixth, it is used to express personal stance, feelings or opinions. For example, *well...*I cannot bear it.

Seventh, it is used to make a request or to make a new start of the conversation. For example, *well...*could we start again…

2.6.6 Functions of *you know*

When Landgrebe (2012) studied this discourse marker from the diachronic perspective, it seemed that *you know* was the discourse marker along the history of language change from its original NP structure as the relationship between a subject (a pronoun) and the predicate started with the verb "know".

It was found that there were six textual functions and six interpersonal functions based on the previous research (Brinton, 1996; Landgrebe, 2012; Müller, 2005; Polat, 2011; Vanda & Peter, 2011), as illustrated in Table 5.

Table 5: Pragmatic functions of the spoken discourse marker you know based on findings of previous research

You know **As a discourse marker**	**Textual functions**	Searching for the right words
		A repair
		More explanation/information
		Leading a quote
		Approximation
		Topic shift
	Interpersonal functions	An implicature of shared or common knowledge
		To acquire an understanding/an involvement from the hearer
		Leading to an imaginary scene
		An implicature of personal opinion/feeling/stance
		Emphasis/repetition
		A hedge/mitigator

The six textual functions are listed as follows.

First, it is used to search for the right words. For example, as they started…and changed the original way…and *you know*…umm…they got something new.

Second, it marks self-repair. For example, it's not like that…*you know*…not like I did that.

Third, it marks more information or explanation. For example, don't be so cruel on them…*you know*…being cruel is just the way showing you are scared of them actually.

80

Fourth, it leads a quotation. For example, she said…*you know…*shall we dance.

Fifth, it marks an approximation. It shows that the following utterance is not the exact information or an incomplete piece of information. For example, you can just laugh…*you know…*like that.

Sixth, it marks a topic shit. For example, A: and that's -- I think that people deserve that, deserve the truth. But some people can't handle the truth. B: *You know,* Frank Sinatra said about writers like that, once he told me…

The six interpersonal functions are listed as follows.

First, it marks an implicature of shared knowledge or common understanding. It is used when the speaker indicates some shared or common knowledge that both the speaker and the hearer have known already. The common knowledge is usually known worldwide. For example, *you know…*people are sometimes mean.

Second, it is used to acquire an understanding from the hearer or to get an involvement from the hearer. For example, they did good too, *you know.*

Third, it leads to an imaginary scene. For example, just kicked it hard…*you know…*when that tall guy ran towards him. and no one could stop him. and he just used the right leg. and kicked him right on his knee so hard…

Fourth, it is used to show the speaker's stance, feelings or opinions. For example, *you know…*I just cannot imagine that.

Fifth, it marks emphasis or repetition. It is used when the speaker uses a strong tone or repeats some lexical items. For example, Oh my God! *You know…*what is that!

Sixth, it is used as a hedge or a mitigator. For example, A: you don't even know the song. B: *you know…*I heard it before…just forgot the name.

2.6.7 Functions of *I think*

It was found that this discourse marker had one textual function and two interpersonal functions (Dehe & Wichmann, 2010; Diessel & Tomasello, 2001), as illustrated in Table 6.

Table 6: Pragmatic functions of the spoken discourse marker I think based on findings of previous research

I think As a discourse marker	Textual functions	Searching for the right words
	Interpersonal functions	An implicature of personal opinion/feeling/stance
		A hedge/mitigator

Its textual function is to search for the right words. For example, *I think…*umm…the whole scenario is too unrealistic…

The two interpersonal functions are listed as follows.

First, it is used to show personal stance, feelings or opinions. For example, …*I think* you are right.

Second, it is used as a hedge or a mitigator. For example, *I think* you probably should leave now.

2.6.8 Functions of *I mean*

It was found that there were four textual functions and four interpersonal functions of the discourse marker *I mean* based on the previous research (Erman, 1987; Fox & Schrock, 2002; Vanda & Peter, 2011), as illustrated in Table 7.

Table 7: Pragmatic functions of the spoken discourse marker I mean based on findings of previous research

		Marking a repair

I mean **As a** **discourse** **marker**	**Textual** **functions**	More explanation/information/exemplification
		Marking topic shift
		Searching for the right words
	Interpersonal **functions**	An implicature of the cause subjectively
		An implicature of personal opinion/feeling/stance
		Emphasis/repetition
		A hedge/mitigator

The four textual functions are listed as follows.

First, it marks self-repair. For example, she works in a pub, *I mean,* a nightclub.

Second, it marks more information or exemplification. For example, it seems not fair, *I mean,* to her.

Third, it leads a topic shift. For example, oh…it's better you check that out…*I mean,* I once rode it and I just felt strange…

Fourth, it is used to search for the right words. For example, it's like…*I mean..*is it gonna be dangerous…

The four interpersonal functions are listed as follows.

First, it marks an implicature of the cause of the prior utterance subjectively. It is used when the speaker tries to explain a prior utterance with his personal reasons. For example, don't you think it's time you put that thing away. *I mean…* look at it, it's antique. you could hurt yourself with it.

Second, it marks personal stance, feelings or opinions. For example, I suggest not talking to him. *I mean,* he's not as good as you think.

Third, it marks an emphasis or repetition. For example, but he is rich! Really rich, *I mean.*

Fourth, it is used as a hedge or a mitigator. For example, *I mean…*could I have a moment alone please.

2.7 Conclusions on reviewing discourse markers

With several secondary studies on discourse markers (Bolden, 2015; Der, 2010; Yang, 2011), this research will have a brief conclusion of the review of the previous research on discourse markers.

Definition and terminology were two main issues in this area. Most of the previous research used the term "discourse markers" (Bolden, 2015). However, until nowadays, different researchers still held onto the agreement of their own. Some researchers (Redeker, 1990) thought that discourse markers would include pragmatic markers, while other researchers (Brinton, 1996, 2008; Fraser, 1993,1999) thought the discourse markers would be one type of marker in pragmatic markers. This research prefers to treat discourse markers as an independent concept in the linguistic area.

Besides, discourse markers were syntactic independence, which meant that they were independent lexical phrases outside of the main semantic sentences. Their major significant role was the different pragmatic functions in utterances, which was also considered the most important feature of discourse markers.

Discourse markers are multi-functional. Researchers used multi-functionality to express them as one of their major features. Since Brinton's (2008) framework of the pragmatic functions of the spoken discourse markers is used in this research, it is believed that the functions of spoken discourse markers can be basically separated into textual functions and interpersonal functions. Moreover, each spoken discourse marker owns its unique multiple functions.

Previous studies indicated that more complicated formations of discourse markers should be studied continuously and tried to ensure their domain of them in terminology.

Previous studies also suggested that in the future, researchers were supposed to go back to the fundamental basis, tried to figure out the root problem, and went further to develop discourse markers and related theories.

2.8 Learner Corpora

Since this research uses learner corpus to do the analysis, in this part, learner corpora are reviewed in general.

Learner corpora were developed with the pace of the development of huge corpora and corpus linguistics. Granger (1998a) was one of the top professionals studying the learner corpora. The following literature review will basically focus on the design of learner corpora, studies on EFL learners by using learner corpora and technical issues in learner corpora.

2.8.1 Design of learner corpora

Based on the early research (Granger, 1998b), learner corpora were considered a new methodology to study language learners' dynamic patterns of their language acquisition.

Based on (Granger, 1998b), features of learner corpora could be divided into the language-related feature, the task-related feature and the learner-related feature. Learner corpora had plenty of issues that should be considered from the beginning of building them until the final analysis based on the data shown in them. From the language-related point of view, learner corpora had written and spoken modes. Based on corpus linguistics, other modes could also be built, such as body gesture corpus, etc. Genres of learner corpora came into mind next. The learner corpora were a large collection of letters, diaries, and argumentative essays or fiction written by students. In each genre, if the styles or topics were not the same, the genres were also different. Thus, at the beginning of building the learner corpora, researchers should strictly

follow the rules and objectives of the proposal to build the corpora for analysis. As Sinclair (2005) mentioned, the contents of a corpus should be selected without regard for the language they contain, but according to their communicative function. Sinclair (2005) also proposed that "corpus builders should strive to make their corpus as representative as possible of the language from which it is chosen" (p. 2). In the task-related feature, the ways of collecting data could be cross-sectional or longitudinal. In the learner-related feature, details of learners, which were also known as metadata should be considered carefully. In the procedure of building the learner corpora, every step should have their consideration in detail. The header information, level of transcription and level of annotation were all on the list of considerations.

Continuing the exploration of building learner corpora, Granger (2004) discussed learner corpora's features from the perspective of some trending controversial issues about the nature of learner corpora. The first controversial issue of a learner corpus is how big it should be. The size of learner corpora will be different based on the original data that could be collected from the researchers' point of view. Also, around the world, learner corpora in large sizes can be found in different areas with different metadata. Learner language is highly variable. Plenty of reasons may have this issue on the variability of learner corpora, such as situational factors, metadata backgrounds or the purposes of learner corpora. Certain large learner corpora may have different genres or sections. Each genre and section may have different variability. Thus, this feature is quite fuzzy and complicated. Another feature is automation. It was illustrated that researchers had the choice of doing automated approaches within four facets: count, sort, compare and annotate with the innovation of technology (Granger, 2004). By using technical tools, i.e., WordSmith Tools (Scott, 1996), as Granger suggested, some results of learner corpora were achieved automatically, including type/token ratio (count), learners' lexico-

grammatical patterning of words (sort), a comparison of native language and learner language (compare), and POS-tagging (annotate).

After the issues on building the learner corpora, Granger (2013) explored the learner corpora's typology. Two trending types have been discussed. The first trending type was contrastive interlanguage analysis (CIA). It was usually considered as the comparison between native speakers and non-native speakers with the difference in using the same language. Since learner corpora were in consideration with the different usage of their second language, the research results should have a huge influence on researchers and learners. Based on the research results, researchers and learners would know how learners used the second language they were learning and what patterns of learners' usage of the second language. The second trending type was computer-aided error analysis (EA). This type had huge differences from CIA because this type involved computer-aided procedures and the whole process was completed by computers with the manipulation of the researchers. Under these circumstances, setting up a standard was an issue. Software tools were also a necessity to be considered. The usual tools were POS-tagging programs and error-tagging programs.

There have been plenty of huge learner corpora around the world through the research from the contributions of different researchers and organizations for many years. Tono (2003) listed several learner corpora that have been developed well around the world, including ICLE, LLC, the UAM corpus, and so forth. Pravec (2002) also listed large learner corpora in this world and introduced them shortly. Although many learner corpora have emerged, they differed from each other in different aspects. Their annotation availability, their data and metadata, and their comparison availability were the elements that made them unique.

The design of the learner corpus in this research is based on the research questions and the objectives of the research. Given that the main purpose of conducting this research is to investigate the spoken discourse markers in English conversation performed by Thai EFL learners, it is therefore reasonable for the researcher to build the learner corpus in spoken mode as the genre of daily English conversation. The learner corpus combines two levels of Thai EFL learners: Thai B1-level EFL learners and Thai C1-level EFL learners. The metadata of all the participants is recorded manually by the researcher. The spoken data is recorded and transcribed into the written data manually by the researcher. The annotation of each spoken discourse marker is examined by the researcher and an interrater based on the criteria illustrated in the Methodology part.

2.8.2 Studies on EFL learners by using learner corpora

By using learner corpora, researchers were able to do studies on EFL learners to discover their use of English. It was believed that learner corpora provided a new methodology to examine the use of English by EFL learners as a whole picture. Among the research that used learner corpora as a methodology, most of them focused on exploring lexical items (Callies, 2015; Granger, 2018; Neff et al., 2004), as reviewed below one after another.

Callies (2015) used learner corpora to explore the use of English by EFL learners in the area of language testing and assessment. Three methodological challenges were discovered in this research. First, to researchers, collecting data to build a corpus was a huge challenge. Second, since building a large corpus required large data from EFL learners, the criteria for selecting the learners were also arguable. The third challenge came from proficiency level as a fuzzy variable in learner corpus in compilation and analysis. This problem was also aligned with the second problem, and the three

challenges were in the same range and degree. Thus, researchers should think of these challenges when they build the learner corpora for analysis.

As one of the experimental studies on SPICLE corpus, Neff et al. (2004) explored the writer stance in learner corpus and native English speaker corpus from the perspective of interactions between readers and writers. The SPICLE was a huge project on L2 English learner corpus from the advanced EFL learners in different backgrounds and native language speakers. By using this learner corpus, previous research has been done to examine L2 English writing from various perspectives. This study was one of the studies that examined the advanced EFL learners' formulaic expressions of expressing writers' stance with the purpose of interacting with the readers. The study used Oxford University Press Wordsmith Tools to have the frequency list and had two major formulaic expressions on the interactions with readers. The first one was *it is + (adverb) adjective + that,* and the other one was specific adjectives that involved probability, obviousness and appropriateness. The results showed that when the adjectives with and without a pre-modifying adverb and the agentless passive were counted together, it seemed that EFL learners had a tendency to overuse the adjectives and the agentless passive constructions in their argumentative essays. Meanwhile, EFL learners tended to use more contrastive conjunctions with different lexical items in the same sense, such as *however, yet, but, nevertheless*, etc. Compared to native English speakers, EFL learners overused these two types of expressions for the purpose of interacting with readers to express their stance. It can be seen that EFL learners had a limitation of using various expressions to have their stance in the writing. Thus, they overused certain expressions to illustrate the same stance of their own.

Granger (2018) used learner corpora to explore two difficult objectives to do the research: collocations and lexical bundles. The research investigated the patterns and

89

frequency of these two types of formulaic expression. In this research, learner corpora presented the frequency list of formulaic expressions used by language learners easily. With the annotation and error tagging procedure, the accuracy and appropriacy of formulaic expression could also be easily known with the analysis. Furthermore, the development of certain formulaic expressions could also be studied with the development of the learner corpora. For lexical bundles, the studies should have a clear definition in order to have criteria for researchers to extract lexical bundles from the learner corpora.

By reviewing some previous studies that used learner corpora as a method, studies can figure out the patterns in the use of English lexical items by EFL learners. By providing machine-readable data, building learner corpora makes it possible for researchers to examine EFL learners' use of English in a systematic way within huge quantitative data.

This research builds the learner corpus to examine how Thai EFL learners perform spoken discourse markers in English conversation. It conducts the experimental study to compare the use of English spoken discourse markers between Thai EFL learners with two English levels and native English speakers. The learner corpus built in this research helps the researcher to examine the differences mentioned above in both quantitatively and qualitatively.

2.8.3 Technical issues in learner corpora

During the development of learner corpora, as well as the studies that used learner corpora to examine the errors of EFL learners, one of the major issues was the technical issue including the development of tagging system, keywords in context searching, development of concordancers and error tagging system, etc. The innovation of technology was necessary for building learner corpora so that future studies would have the data and criteria to continue to use learner corpora for analysis.

For many years, researchers have explored POS tagging by using computers (Leech et al., 1994). POS tagging has also been developed several times. In the meantime, researchers also tried to develop the discourse markers tagging by using computer programs (Heeman et al., 1998; Popescu & Zufferey, 2010).

Heeman et al. (1998) were in one of the early teams that started to use corpora to do POS tagging altogether with the tagging of discourse markers. Meanwhile, they also tried to discover the features of discourse markers after they were able to have the POS tagging by the computer program. The study used Trains Corpus (Allen et al., 1995) and used symbols and techniques to start the POS tagging for discourse markers. The research found that by using Markov model approach, they were able to identify discourse markers in spoken dialogue. However, it could not be accomplished automatically. The process of the identification was considered one part of the POS tagging procedure. If this technique could keep developing, tagging discourse markers would be included into the POS tagging system. Thus, it would reduce the errors of POS in spoken language. The process of POS tagging discourse markers was not considering discourse markers as single phrases or lexical items. They tried to interpret them as acoustic signals. Therefore, the process and the idea of POS tagging discourse markers would be more complicated than the original POS tagging words. The research indicated that if researchers could automatically identify discourse markers in the process of POS tagging, it would help hearers to predict the following speech.

Popescu and Zufferey (2010) tried to use computer programs to mark discourse markers: *like* and *well*. The study chose the sociolinguistic feature for the marking because it was the most frequent function of the discourse marker used by speakers in utterances. The research specifically chose two discourse markers *like* and *well* because they were identified as discourse markers in the previous research and they

were in normal lexical classifications in the traditional linguistic point of view. Thus, when researchers entered them into the computer system, with three selected features, a computer was supposed to select both words and classified them into two categories: discourse markers and non-discourse markers. The ICSI-MR corpus was used for analysis. The results of this computational procedure showed that lexical collocations were the easiest and first choice for computers to select discourse markers when the computer was asked to search them. Computers could search both discourse markers with their roles, occurrences and positions in utterances. The prosodic tagging helped computers to separate discourse markers and non-discourse markers, but it must be noted that it was not the only option for a computer to choose. With the help of sociolinguistic features, researchers could examine the result in detail and see whether there would be mistakes or not.

Therefore, it can be seen tagging discourse markers is not similar to POS tagging. The POS tagging follows part-of-speech models while tagging discourse markers needs more complicated criteria. Moreover, automatic tagging discourse markers may cause incorrectness. It was thus suggested that tagging the discourse markers manually may be more appropriate for the studies on them.

Another technical issue related to learner corpora is error tagging. Granger (2003) explored learner corpora in CALL. Before learner corpora could be introduced into the CALL system, there was one important thing to do for learner corpora: error tagging. It was also aligned with the findings of Meurers (2012) that error tagging learner corpora was seen as one of the significant steps for building learner corpora with different purposes. For example, if the purpose of building a learner corpus was to examine the error use of verbs in EFL learners, error-tagging would be on the point of verbs in the learner corpora.

With the development of the technology, computer-aided error analysis was done in certain areas of error tagging. The analysis of the error in words, lexical items, lexico-grammar has been done by using the error-tagging program. Different studies have tried to do the error tagging from different perspectives. For example, Ramos et al. (2010) explored collocation error tagging in learner corpora and a theoretical way of doing it.

To sum up, different researchers still intended to make learner corpora into a systematic methodology and linguistic sub-category. Future studies should focus more on the annotation in learner corpora with the development of technology and software.

The annotation of the learner corpus in this research includes two parts: the POS-tagging automatically by the instrument CLAWS and the identification of spoken discourse markers manually. Based on the criteria for the identification of spoken discourse markers presented in the Methodology part, the researcher tags all the spoken discourse markers that meet the criteria.

2.8.4 Conclusions on learner corpora

This research uses Gilquin's (2005) study on learner corpora to have a brief conclusion of the review on learner corpora.

Learner corpora had a huge connection with the study of the use of English by EFL learners. Their relationship bounds each other tightly. As definitions of learner corpora were also in a complicated way to deliver, this research uses the definition of learner corpora proposed by McEnery et al. (2006) that learner corpora should be "collection of machine-readable authentic texts (including transcripts of spoken data)" (p. 5) which is sampled to be representative of learners' interlanguage. What made learner corpora special was that the original data was collected from the language learners and in many researchers' eyes, their language could be called the inter-

language, which was worth doing analysis. The representativeness of the data in the learner corpora is treated as the language patterns from the language learners themselves. Thus, the issue traces back to the beginning that learner corpora were seen as the huge collection of original language data directly from learners. Therefore, it should have an enormous contribution to the development of SLA and ELT. The degree of naturalness in the learner corpora has been examined in highly perspectives since the researchers had specific objectives and purposes for collecting the data and building the corpora. In Krashen's (1977) sense in language study, the topic that was related to the concept of naturalness was also the editing of data in the corpora. Therefore, before considering building a learner corpus, researchers should have thoughts on the purpose of building the corpora, what aspects they needed to analyze, and how they would use the learner corpora to accomplish the objectives of their thoughts. Another important issue before building the learner corpora is the typology issue. Since researchers intended to use English produced by EFL learners as the original metadata, their language production can be separated into various types. This issue has been discussed a lot in the previous reviews. The data should be either the written form or the spoken form. The metadata, which shows the information of the participants in the corpora, should be at the same level or cross-section.

When it is related to the design and the procedure of building the learner corpora, environment, tasks and learner variables should be considered carefully. The procedure of collecting data may be a long and complicated procedure for researchers. As for wishing to collect valid data based on the objectives of the study, researchers need to have the corresponding environment to have the data original and real. Selecting the participants is one of the complicated issues when researchers collect data. Thus, tasks should be told appropriately to participants in order to collect the original data in real situations. On the whole, all the problems seem to be related to

94

the issue of learner variables based on the findings of the previous research (Gut, 2012; Jendryczka-Wierszycka, 2009).

The final issue of learner corpora is related to the limitation of the previous studies. First, although there have been several large learner corpora developed by different researchers from different regions, it is still a striking fact that the development of learner corpora was not in balance, in which two levels of imbalance were found. The first imbalance is that it seemed that learner corpora were still only developed in certain areas. Many regions in the world were not in the procedure of building learner corpora. The second imbalance is that typology in learner corpora was not in the same quantity. The spoken learner corpora should be further developed as well as written data. Other issues were also considered. Limited access to learner corpora would be a huge issue to push corpus linguistics into a more practical way. Some of the data could not be shared with every researcher.

Meanwhile, learner corpora have contributed to the development of second language acquisition (SLA) (Granger, 2003, 2004, 2013). Technical issues have also been found in this area (Myles, 2005). Furthermore, the studies on building learner corpora have been in different areas to help researchers examine the patterns performed by EFL learners, i.e., using learner corpora to study EAP (Gilquin et al., 2007). Due to the little correlation to this research, the literature review was not focused on these areas.

Chapter 3

Methodology

This research combines both quantitative and qualitative methods to investigate the use of the English spoken discourse markers by Thai EFL learners.

This chapter introduces the population and samples of this research, the instruments used in this research, data collection, data analysis and the whole procedure of the research.

3.1 Population and samples

In the previous studies of the comparison between native English speakers and non-native English speakers, researchers studied different EFL learners with different English levels (Buysse, 2012; Diskin, 2017; Fung & Carter, 2007; House, 2013; Polat, 2011). Intermediate-level (B1-level) learners and advanced-level learners have been studied the most (Bax, Nakatsuhara & Waller, 2019; Sitthirak, 2013).

Moreover, different previous research had different numbers of participants. Hellermann and Vergun (2006) had 17 adult learners to be the participants, which was the minimum number that this research could find out, except that Polat (2011) had a case study with only one participant. Asik and Cephe (2013) had 20 participants involved in the research. Tagliamonte (2005) involved 26 participants. Diskin (2017) had 42 participants with both Irish and Chinese EFL learners. Fung and Carter (2007) collected data from 49 participants. Aijmer (2004) had 50 interviews. Trillo (2002) chose 54 samples from the CHILDES corpus. Therefore, it can be seen that the number of participants was in the range of 20-60 participants.

This research focused on L1 Thai EFL learners who were raised and educated in Thailand to the undergraduate level. They were mainly exposed to the Thai language environment where English was not often used and was regarded as a foreign

language. According to Wudthayagorn (2018), L1 Thai undergraduate students should reach the English level at B1 level based on the CEFR assessment standard (Council of Europe, 2001). To examine the differences in the use of the English spoken discourse markers, this research focuses on Thai B1-level EFL learners and Thai C1-level EFL learners.

Meanwhile, this research had 30 Thai B1-level EFL learners and 30 Thai C1-level EFL learners from Chulalongkorn University to be the participants, accounting for a total of 60 Thai participants involved in this research. The number of 30 participants at each English level was selected as because it was considered the minimum number to conduct a test in statistical analysis (Ross, 2017).

To select the participants in each English level, the CU-TEP scores were used as a standard to examine both B1 level and C1 level.

Based on Wudthayagorn (2018), the CU-TEP scores have been mapped with the standards in CEFR. According to the research, CU-TEP has the cut-off score ranges with the equality of CEFR levels from A2 to C1, including B1 and C1. The full score of CU-TEP is 120 points. The cut-off points of B1-level are 35-69 points, while the cut-off points of C1 level are 99-120 points. Thus, the research selected qualified participants within each cut-off point randomly to let them participate in this study.

3.2 Research instruments

The research instruments in this study were constructed based on research questions and the research design. Seven instruments were used in this research. They were face-to-face prepared questions on selected topics; SBCSAE; the recording machine; CLAN as the transcriber; AntConc; CLAWS 4 and log-likelihood. Each instrument will be introduced one after another in detail. The choice and the basic information of each instrument are illustrated with the research procedure as follows.

3.2.1 Topics and questions for the conversation experiment

First, materials that have been used in the previous studies and participants are introduced generally.

The materials used in previous research can be separated into two types. The first type is the spoken language corpora that have already been built before. Researchers used them to do the analysis directly. For example, Dobrovoljc (2017) directly used GOS corpus, which was the spoken Slovene that had already existed. The GOS corpus contains spontaneous speech. The second type is the materials that researchers collected from learners. In most of the research, the data was in the genre of conversations. There are two types of conversations. The most collected conversations were informal daily conversations. For example, the conversations between teachers and students during consultation time (House, 2013); the daily dialogues (Popescu-Belis & Zufferey, 2010); the daily conversations (Lee-Goldman, 2010; Polat, 2011), etc. There was also spoken data from formal conversations such as formal interviews (Buysse, 2012; Diskin, 2017). Aijmer (2004) always tried to convince that the pragmatic markers were significant and relevant to communication.

Based on the previous studies, it is clear to see that informal daily conversations have been used the most for the experiments to collect the data. Thus, this research collected data from conversations in English and used them for data analysis.

In order to control the fairness of each conversation with each participant, the researcher decided to choose topics and prepared the same questions when starting a conversation on each topic with the participants.

Choosing topics is an important step. First, the research was intended to choose the topics that related to participants' life. Second, the topics should be in a normal sense without touching sensitive points, such as political issues.

The researcher first had a collection of five contemporary commercial textbooks for teaching or learning oral communication in English and selected the top five topics among these textbooks for conversations. The researcher mainly chose the textbooks published within the latest ten years and published by the worldwide major publication corporations or presses. The five commercial textbooks were as follows:

1. Speak Out (Eales & Oakes, 2011)

2. Impact Topics (Day & Yamanaka, 2003)

3. Dynamic Listening and Speaking (Jeon & Perderson, 2009)

4. Speaking B1+Intermediate (Pelteret, 2012)

5. Pathways 2 (Chase & Johannsen, 2012)

These textbooks are all for EFL learners to improve their English-speaking skills. After calculating the frequency of each topic among five textbooks, the top five topics are as follows:

1. Education (school and university life)

2. People (meeting people, describing people and social interaction)

3. Problems and solutions (problems in society, giving opinions, delivering suggestions and solving problems)

4. Relationship (family, neighbors and friends)

5. Travel (trips and describing countries)

In order to fully prepare the conversation, each topic has around five questions chosen from these textbooks. Each participant had a conversation on all five topics. However, it must be noted here that the daily conversation is natural and uncontrollable due to different answers. It was thus conceivable that with each participant's different answers, certain questions were not asked and other questions related to the answers from the participants were illustrated in the real situation. Here are some sample questions on each topic and please see the Appendix also.

Questions relevant to the topic of "education":

1. Tell me something about your school life before university. What was it like?

2. What was your senior high school life like?

3. What do you think of your university life so far? How is it different from your school life?

Questions relevant to the topic of "people":

1. How do you describe your personality?

2. Have you kept in touch with your friends in primary school or high school?

3. How do you usually make friends?

Questions relevant to the topic of "problems and solutions":

1. Was there anything that bothered you when you were growing up?

2. How did you solve the problem when you were bothered by the things you just mentioned when you were growing up?

3. Has anything bothered you in your university life so far?

Questions relevant to the topic of "family":

1. What was your childhood like?

2. Do you think the relationship between you and your parents is very close?

3. What is the best thing about your family?

Questions relevant to the topic of "travel":

1. Do you like traveling? Why?

2. Tell me one place that you visited before.

3. What do you usually do during traveling?

3.2.2 Selection of native English speaker corpus

One of the important issues in this research was to choose a native English speaker corpus that was suitable to make comparisons and matched with the learner corpora built in this research. After comparing different native English speaker corpora in the spoken mode, i.e., CANCODE, BNC, etc., with consideration of the feasibility, it seems that Santa Barbara Corpus of Spoken American English (henceforth SBCSAE) was a suitable corpus as the native English speaker corpus in this research.

SBCSAE is available and free to download certain materials on the website: https://www.linguistics.ucsb.edu/research/santa-barbara-corpus. According to the introduction of the official website of this corpus, it includes approximately 249,000 words, with audio and corresponding transcription. The four parts of the project are all free to download to anyone with all the files of audio, transcription and metadata. Therefore, it is assumed that there should be no ethical problem as long as the research quotes it appropriately.

SBCSAE is one part of International Corpus of English (ICE). It is a corpus that includes all naturally recorded English spoken interactions across different places in the whole area of the United States of America. It includes different genders, ages, regions and social backgrounds. There are varieties of recording the spoken interaction. The face-to-face conversations have been the most common type in daily life. There are also other types including phone conversations, on-the-job talk, card games, food preparations and more. From this short introduction of SBCSAE, three points can be concluded. First, it is a comparatively large corpus that includes enormous data. It covers large areas of the USA so that the data has high representativeness. Second, it is a spoken corpus that represents English spoken data from the American English native speakers. Third, the spoken data mainly came from daily conversations. Thus, the spoken type would not be academic but informal daily

conversation representativeness. One of the examples of the transcription in CHAT format is shown in Figure 4.

Figure 4: An extract of transcription from SBCSAE in CHAT format

(Du Bois et al., 2000-2005)

```
@Begin
@Languages:    eng
@Participants: ALAN ALAN Speaker, JON JON Speaker
@Options:      CA, caps, bullets
@ID:    eng|SBCSAE|ALAN|||||Speaker|||
@ID:    eng|SBCSAE|JON|||||Speaker|||
@Media: 60, audio
@Comment:      Shaggy Dog Story
@Comment:      Face-to-face casual conversation recorded in an office in
        Shreveport, Louisiana. The two speakers, Jon (age 72) and Alan (age
        66) are friends/co-workers taking a break from work. Alan is
        primarily telling Jon about his travel adventures and interests
*ALAN:  (..) &=in I got a story to tell you •
        it's a shaggy dog story •
        but as long as we're talking about that recorder and six-hundred dollars •
        &=in &=lengthened I'll tell you the story •
        &=in &=lengthened oh gosh •
        uh &=ex •
        Rae and I and Sue and Buddy •
        (..) took a trip •
        (..) to Mexico City . •
*JON:   (..) &=SNIFF . •
*ALAN:  It must've been •
        (..) four to six months after my dad died . •
*ALAN:  That's how I remember it •
        he ⌈ died in slixty-s:- +... •
*JON:        ⌊ Oh God ⌋ . •
*ALAN:  (..) December sixty-seven •
        so •
        &=in &=lengthened sometime in sixty-eight we took this trip •
        we'd been (..) talking about it for a while •
        (..) uh: •
        flew down to Mexico City •
        (..) uh we •
        &=ex c- think of the name of my hotel •
        which wouldn't mean anything now •
        but we ended up in a (..) fabulous hotel •
        (..) uh: •
        (..) first night •
        we were &{l=VOX very unhappy &}l=VOX with our rooms •
        we got down there •
        &=in &=lengthened and the next morning •
        Buddy •
        who's a (..) early riser anyhow •
        was probably up (..) four o'clock •
        and he went down there complaining to the manager •
        (..) So •
        (.) cause it was not w- the accommodation we were supposed to have had •
        we checked in about eight o'clock at night or so •
        which is •
```

In this research, the spoken data was collected from the conversations illustrated by Thai EFL learners. Although as the requirement of collecting the data, the conversation took place in a specific setting with selected topics and locations, the conversations were informal, including changing ideas, sharing experiences, discussing problems and issues, and so forth. The experimental procedure is similar to the daily conversations between two people talking in English only. Thus, SBCSAE is

the most suitable corpus as the native speaker corpus to have the comparison in this research.

Since SBCSAE is available to researchers, it has been widely used in previous research (Aijmer, 2013; Polat, 2011; Sakita, 2013) as a native English speaker corpus to be compared with the spoken data from different perspectives. Polat (2011) used SBCSAE to make the comparison of the use of discourse markers between non-native English speakers and native English speakers. Aijmer (2013) used SBCSAE as one of the native English speaker corpora to have an explanation of pragmatic discourses. Sakita (2013) also used SBCSAE to study the discourse marker *well* in the spoken data. Therefore, this research followed the step of previous studies and chose SBCSAE as the native English speaker corpus.

3.2.3 Spoken data collection instruments

In order to collect spoken data from the participants, two instruments were used in the research.

First, an iPhone recorder was used to record students' audio in mp3 format. Each participant was informed that the conversation between the researcher and the participants would be recorded before the recording, including that the recording is only for research, not for certain tests and will never influence their grades in the university; or their personal information will never be given out to the public, and so forth.

Second, all the spoken data were transcribed into written form by using CHAT (Codes for the Human Analysis of Transcripts). This software was designed and innovated by MacWhinney (2000) from the 1990s until 2000. It was originally for the research of exploring children's language acquisition and for building up a corpus that contained the early language patterns of children. It can be downloaded on the website with the address: https://childes.talkbank.org. It can also identify errors and

mark them with different signs. The correctness of transferring is above 98% (MacWhinney, 2000). One of the written texts transferred from the original oral data is presented in Figure 5.

Figure 5: An example of transcription from CHILDES Project

(MacWhinney, 2000)

```
@Begin
@Languages:        eng
@Participants: CHI Ross Child, FAT Brian Father
@ID:        eng|macwhinney|CHI|2;10.10|||Target_Child|||
@ID:        eng|macwhinney|FAT|35;2.|||||Target_Child|||
*ROS:       why isn't Mommy coming?
%com:       Mother usually picks Ross up around 4 PM.
*FAT:       don't worry.
*FAT:       she'll be here soon.
*CHI:       good.
@End
```

3.2.4 Instruments for analysis

The research examined discourse markers from quantitative and qualitative perspectives. Thus, based on the research questions, three instruments were used.

First, after collecting all the data and transcribing them into written text, all the words in the data were tagged by their part-of-speech (POS). To tag POS of each word, the software CLAWS4 was used.

The basic information about CLAWS4 can be found on the website: http://ucrel.lancs.ac.uk/claws/. Illustrated in the previous parts, POS tagging is the commonest form of corpus annotation. It was first established and developed by UCREL at Lancaster. The system has been continuously developed since the early 1908s. CLAWS4 was used to tag POS for British National Corpus (BNC) (Leech, Garside & Bryant, 1994). CLAWS has consistently achieved 96-97% accuracy (the precise degree of accuracy varies according to the type of text). Judged in terms of

major categories, the system has an error rate of only 1.5%, with c.3.3% ambiguities unresolved, within the BNC. CLAWS4 has also been used in several studies to tag learner corpora. Granger (1998) used CLAWS to POS tag learner corpora. It was encouraged by Granger (1998) that researchers may use POS tagging for learner corpora. Granger (2013) illustrated that POS-tagged learner corpora allowed more refined linguistic analysis (p. 28). Some huge learner corpora also used CLAWS to POS tag, such as the PELCRA learner corpus, the HKUST learner corpus, the JEFLL learner corpus, etc. Granger (2013) also explained that the researchers should examine the POS tagging done by CLAWS, for no automatic tagging was 100% right, and CLAWS was designed to tag native English speaker corpora in the first place.

When entering the website: http://ucrel-api.lancaster.ac.uk/claws/free.html, researchers can put in texts and CLAWS will tag POS of the words in the text. If researchers do not know how to input or they have questions about the format of the input, they can click on the website: http://ucrel.lancs.ac.uk/claws/format.html. One of the short examples of using CLAWS4 with the input sentence and POS output is in Figure 6 from the official website.

Figure 6: An example of POS tagging by using CLAWS 4

(Leech et al., 1994)


```
<text>
The quick brown fox jumps over the lazy dog.
</text>

CLAWS outputs a verticalised form of the text where each word has a list of possible POS tags. The most likely tag is the first in the list.

0000001 001 **6;0;START                         01 NULL
0000001 002 ------------------------------------------------
0000003 010 The                                 93 AT
0000003 020 quick                               93 [JJ/99] RR@/1 NN1%/0
0000003 030 brown                               93 [JJ/93] NN1@/7 VV0%/0
0000003 040 fox                                 93 [NN1/100] VV0@/0
0000003 050 jumps                               93 [VVZ/97] NN2@/3
0000003 060 over                                93 [II/59] RP/41 NN1%/0 JJ%/0
0000003 070 the                                 93 AT
0000003 080 lazy                                93 JJ
0000003 090 dog                                 93 [NN1/100] VV0%/0
0000003 091 .                                   03 .
0000004 001 **7;7;text                          01 NULL
```

As illustrated above, CLAWS4 still has the possibility of tagging in error. The researchers need to examine the tagging after CLAWS4 does its work.

In this research, CLAWS4 was used to tag POS of each token because POS of each token was used to identify the spoken discourse markers and non-spoken discourse markers in the process.

Second, the research needed to search for the discourse markers to examine their functions. Thus, the research used AntConc 3.5.8 (Macintosh OS X) 2020 to show the discourse markers in concordancers.

Different versions of AntConc software for Windows and Macintosh systems can be found on the website as follows: http://www.laurenceanthony.net/software/antconc/. The software was established and developed by Laurence Anthony (Anthony, 2019) who has been working at Waseda University in Japan. The new version now is AntConc 3.5.8 for both Windows and Macintosh systems. One of the major uses of this software is for editing observed corpus and makes the file into concordances as requested in Figure 7.

Figure 7: An example of AntConc interface retrieved from this research
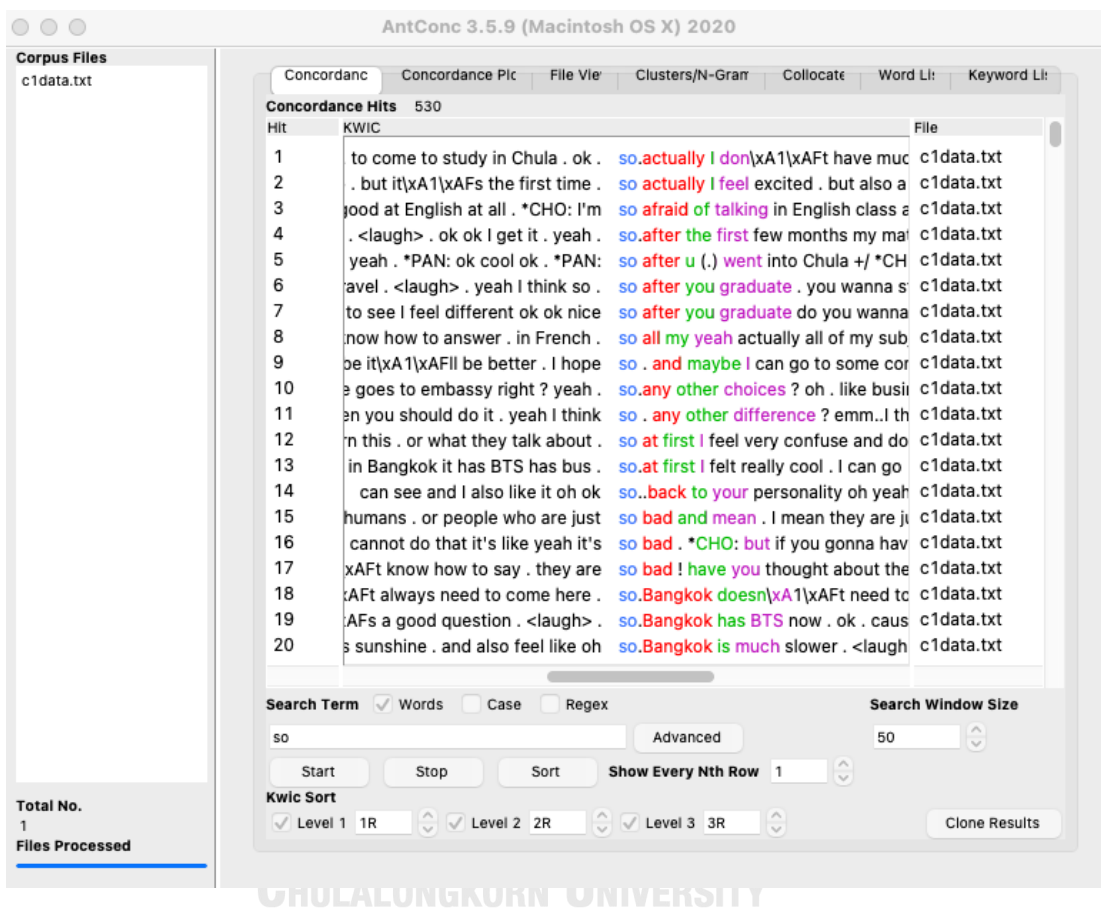
In this research, AntConc was used as two functions. The first function is to calculate the tokens of the corpora. Second, it presented the descriptive data and provided the concordancers for function analysis.

One of the interfaces of searching keywords in AntConc is presented in Figure 8.

Figure 8: An example of searching words in AntConc retrieved from this research
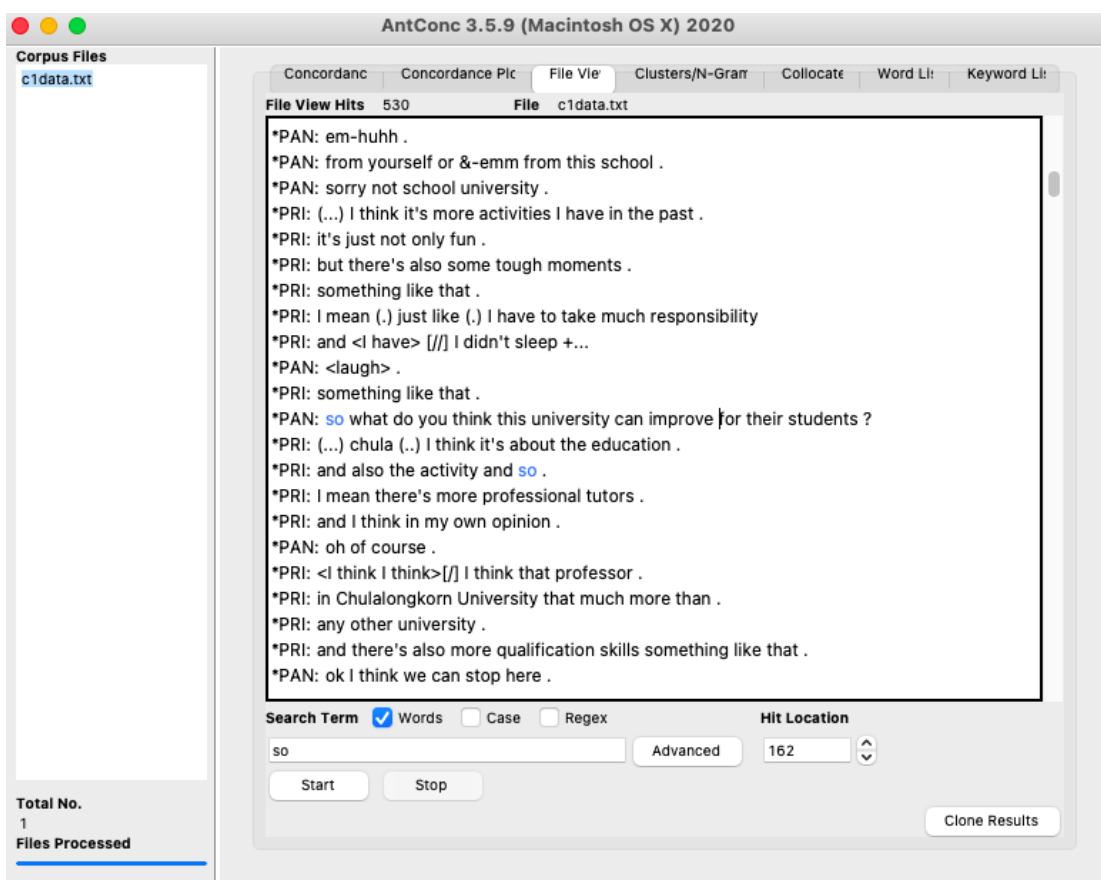
3.2.5 Criteria for tagging the spoken discourse markers

After using AntConc to locate the spoken discourse markers that were under investigation in this research, tagging the spoken discourse markers was the next step.

Tagging discourse markers is a complicated and challenging task, for there has not been a systematic and comprehensive method to tag discourse markers in agreement so far. In the previous literature review, the research conducted by Popescu

and Zufferey (2010) and Heeman et al. (1998) both used different criteria to tag discourse markers. Meanwhile, even though both research used computers to assist the tagging, the researchers pointed out that tagging manually should be involved. Therefore, since there has been no agreement on the issue of tagging discourse markers, and there has no computational system like CLAWS to tag discourse markers, this research tagged the spoken discourse markers under investigation manually with an interrater who was taught the criteria of tagging the spoken discourse markers in this research as follows.

Setting up the criteria for identifying whether a word is a spoken discourse marker is significant. This research identified the spoken discourse markers from two perspectives. The first perspective is excluding the ones whose nature does not belong to a spoken discourse marker. This research used *Oxford Dictionary of English (3rd Edition)* (2010) as a reference to examine the original or traditional meanings of each lexical item. This is because *Oxford Dictionary of English (3rd Edition)* (2010) is currently the largest single-volume English language dictionary published by Oxford University Press (Stevenson, 2010). The second perspective is to identify the features of the spoken discourse markers based on the previous research. The details of the criteria of identifying each spoken discourse marker in this research are presented as follows.

The word: *like*

The following criteria have been found or used in the previous studies (Anderson, 2000, 2001; Diskin, 2017; Müller, 2005; Polat, 2011; Popescu-Belis & Zufferey, 2010; Tagliamonte, 2005).

Criteria of identifying *like* as a non-discourse marker

1. The word *like* is in the position of being a verb.

2. The word *like* is in the position of being a preposition meaning having similar characteristics or qualities.

3. The word *like* is in the position of being a preposition to lead examples for prior utterances in detail.

4. The word *like* is in the position of being a conjunction to have a comparison with two similar items.

5. The word *like* is in the position of being a noun meaning the same or similar interests.

6. The word *like* is in the position of being an adjective meaning the same or the similar.

7. The word *like* is in the position of being an adverb meaning similar.

Criteria of identifying *like* as a discourse marker

1. The word *like* is examined to own one of the pragmatic functions concluded in 2.6.1.

2. The word *like* is between two utterances with no semantic meaning independently.

3. The word *like* is phonologically unstressed.

4. The word *like* is used with more time of pause compared to the speaker's usual oral speed in the same context.

The word: *so*

The following criteria have been found or used in the previous studies (Blakemore, 2014; Buysse, 2012; Müller, 2005; Tagliamonte, 2005).

Criteria of identifying *so* as a non-discourse marker

1. The word *so* is in the position of being a conjunction to lead a result or an effect with a direct cause in the prior utterances.

2. The word *so* is in the position of being an adverb meaning to such a great extent to modify an adjective.

3. The word *so* is in a fixed phrase such as *so to speak, and so on,* etc.

Criteria of identifying *so* as a discourse marker

1. The word *so* is examined to own one of the pragmatic functions concluded in 2.6.2.

2. The word *so* is used with more time of pause compared to the speaker's usual oral speed in the same context.

3. The word *so* is phonologically unstressed.

The word: *well*

The following criteria have been found or used in the previous studies (Blakemore, 2014; Fung & Carter, 2007; Lam, 2009; Müller, 2005; Norrick, 2001; Sakita, 2013).

Criteria of identifying *well* as a non-discourse marker

1. The word *well* is in the position of being an adverb meaning in a good or satisfactory way.

2. The word *well* is in the position of being an adjective meaning in good health.

3. The word *well* is in the position of being a noun.

4. The word *well* is in a fixed phrase such as *as well, be well and good,* etc.

Criteria of identifying *well* as a discourse marker

1. The word *well* is examined to own one of the pragmatic functions concluded in 2.6.3.

2. The word *well* is at the very beginning of an utterance.

3. The word *well* is used with more time of pause before the following utterance is produced compared to the speaker's usual oral speed in the same context.

4. The word *well* is phonologically unstressed.

The expression: *you know*

The following criteria have been found or used in the previous studies (House, 2009; Landgrebe, 2012; Müller, 2005; Polat, 2011; Vanda & Peter, 2011).

Criteria of identifying *you know* as a non-discourse marker

The expression *you know* is in the position of being a noun-verb phrase structure at the beginning of an utterance as the subject-verb form. It can lead to an object or a clause.

Criteria of identifying *you know* as a discourse marker

1. The expression *you know* is examined to own one of the pragmatic functions concluded in 2.6.4.

2. The expression *you know* is used with more time of pause before the following utterance is produced compared to the speaker's usual oral speed in the same context.

3. The expression *you know* is phonologically unstressed.

4. The expression *you know* is in the independent position, unrelated to any utterances from the syntactic point of view. It does not lead to an object or a clause.

The expression: *I think*

The following criteria have been found or used in the previous studies (Aijmer, 1997; Dehe & Wichmann, 2010; Diessel & Tomasello, 2001).

Criteria of identifying *I think* as a non-discourse marker

1. The expression *I think* is in the position of being a noun-verb phrase structure at the beginning of an utterance as the subject-verb form. It can lead to an object or a clause.

2. The expression *I think* is stressed in the utterance.

Criteria of identifying *I think* as a discourse marker

1. The expression *I think* is examined to own one of the pragmatic functions concluded in 2.6.5.

2. The expression *I think* is used with more time of pause before the following utterance is produced compared to the speaker's usual oral speed in the same context.

3. The expression *I think* is phonologically unstressed.

4. The expression *I think* is also produced in the middle of an utterance or at the end of an utterance. The semantic meaning of this utterance does not change at all if *I think* is deleted.

The expression: *I mean*

The following criteria have been found or used in the previous studies (Erman, 1987; Fox et al., 2002; Landgrebe, 2012; Vanda & Peter, 2011).

Criteria of identifying *I mean* as a non-discourse marker

1. The expression *I mean* is in the position of being a noun-verb phrase structure at the beginning of an utterance as the subject-verb form. It can lead to an object or a clause.

2. The expression *I mean* is stressed in the utterance.

Criteria of identifying *I mean* as a discourse marker

1. The expression *I mean* is examined to own one of the pragmatic functions concluded in 2.6.6.

2. The expression *I mean* is used with more time of pause before the following utterance is produced compared to the speaker's usual oral speed in the same context.

3. The expression *I mean* is phonologically unstressed.

4. The expression *I mean* is also produced in the middle of an utterance or at the end of an utterance. The semantic meaning of this utterance does not change at all if *I mean* is deleted.

After the identification of all the discourse markers that should be investigated in this research, the researcher used [DM] to tag each spoken discourse marker in the examples as follows.

…they were like [DM] …don't care about education…

…so [DM]…what it mean?...

…treated me not so good, but, I mean [DM], I know I didn't do that…

3.2.6 Log-likelihood

In this research, a statistical method was needed to test the significant difference in the use of the spoken discourse markers between the learner corpus and the native English speaker corpus.

The learner corpus with the data of Thai EFL learners built in this research was divided into two sub-corpora: Thai B1-level EFL learner corpus and Thai C1-level EFL learner corpus. Three corpora were compared with the native English speaker corpus for statistical and function analysis. The frequency of each spoken discourse marker performed by both Thai B1-level EFL learners and Thai C1-level EFL learners in this research was compared to the one in SBCSAE one after another by using log-likelihood (henceforth LL) test for the significant difference test.

LL has been used to do the significant difference test in comparison of the keywords' frequency in two situations (Rayson et al., 2004; Rayson & Garside, 2000): a comparison between two similar-sized corpora or a comparison between two unbalanced sized corpora.

Rayson and Garside (2000, 2004) introduced LL test and attested the reliability of the LL test when researchers used it to do the significant difference test between

two corpora, where they discovered that the LL test had a high reliability to compare the frequency of lexical items between two unbalanced sized corpora. (Rayson & Garside, 2000; Rayson, Berridge & Francis, 2004). Thus, this research chose the LL test for quantitative analysis.

As was recommended by Rayson et al. (2004), this research set up the p-value at 0.01% level, with the LL critical value at 15.13 to be the cut-off point.

The 2 x 2 table of the LL Wizard created by Rayson was used to do the significant difference test for the comparison of each spoken discourse marker, as shown in the link: http://ucrel.lancs.ac.uk/llwizard.html.

The 2 x 7 table created by Hardie was used to do the significant difference test for the comparison of all the six spoken discourse markers, as shown in the link: http://corpora.lancs.ac.uk/sigtest/#extraHelp.

Table 8 summarizes the instruments used to answer each research question, with the types of statistics corresponding to each research question.

Table 8: Summary of the research instruments

Research questions	Instruments	Statistics
1. Is there a significant difference in use of spoken discourse markers between Thai EFL learners and native English speakers?	1. Topics for conversations	Descriptive statistics
	2. SBCSAE	Descriptive statistics
	3. iPhone Voice Memos	Descriptive statistics
	4. CLAN	Descriptive statistics
	5. CLAWS 4	Descriptive statistics
	6. AntConc 3.5.8	Descriptive statistics
	7. LL	Inferential statistics
2. Is there a significant difference in use of spoken discourse markers between Thai B1-level EFL learners	1. Topics for conversations	Descriptive statistics
	2. iPhone Voice Memos	Descriptive statistics
	3. CLAN	Descriptive statistics
	4. CLAWS 4	Descriptive statistics

and Thai C1-level EFL learners?	5. AntConc 3.5.8	Descriptive statistics
	6. LL	Inferential statistics
3. What are the pragmatic functions of spoken discourse markers used by Thai EFL learners?	AntConc 3.5.8	Descriptive statistics
4. What are the different pragmatic functions of spoken discourse markers used by Thai EFL learners compared to native English speakers?	1. AntConc 3.5.8	Descriptive statistics
	2. SBCSAE	Descriptive statistics

3.3 Data collection

Based on the research questions and the research objectives, the procedure of collecting the data is illustrated as follows.

The researcher prepared the top five topics from the commercial textbooks as follows:

1. Education (school and university life)

2. People (meeting people, describing people and social interaction)

3. Problems and solutions (problems in society, giving opinions, delivering suggestions and solving problems)

4. Relationship (family, neighbors and friends)

5. Travel (trips and describing countries)

In order to fully prepare the conversation questions, each topic had five questions chosen from these textbooks. All five topics were included in the conversation with each participant. However, due to the flexibility of the conversation with each

116

participant, not all the questions were asked. Meanwhile, some questions that were not prepared were asked according to the real situation.

The data collection was conducted from September 2020 to November 2020. A total of 60 participants were involved, with 30 Thai B1-level EFL learners and 30 Thai C1-level EFL learners. Each participant had around 20-minute daily English conversation with the researcher. Each conversation was recorded by Voice Memos Application on iPhone 7 under the circumstances where each participant understood their voice would be recorded.

After finishing all the conversations, the researcher used CLAN to transcribe the spoken data into written data by using the format of CHAT to do the data analysis.

3.4 Data analysis

After the procedure of data collection, the researcher analyzed the data to answer each research question. Before the data analysis, the learner corpora were built.

The learner corpus in this research consisted of the spoken data in its written form from 60 Thai EFL learners, including two sub-corpora. The first sub-corpus is the oral data from 30 Thai B1-level EFL learners. The second sub-corpus is the oral data from 30 Thai C1-level EFL learners.

To answer the first question: is there a significant difference in use of spoken discourse markers between Thai EFL learners and native English speakers, two comparisons were made: the comparison in use of all spoken discourse markers and the comparison in use of each spoken discourse marker. The comparison was done for three pairs of corpora: Thai-SBCSAE, Thai B1-SBCSAE, Thai C1-SBCSAE.

To answer the second question: is there a significant difference in use of spoken discourse markers between Thai B1-level EFL learners and Thai C1-level EFL learners, two steps were processed: the comparison in use of all spoken discourse

117

markers and the comparison in use of each spoken discourse marker between Thai B1-level EFL learners and Thai C1-level EFL learners.

To answer the third question: what are the pragmatic functions of spoken discourse markers used by Thai EFL learners, the researcher used qualitative data analysis based on the concordances provided by the learner corpora to analyze the pragmatic functions of each spoken discourse marker performed by Thai EFL learners. Based on the review of the methodology in previous research, this research used Brinton's (2008) framework of functions of discourse markers to guide the analysis of the functions of spoken discourse markers under investigation, combining with the specific functions of each spoken discourse marker under investigation that have been concluded in the literature review. This is because Brinton's (2008) framework provides the theoretical foundation of the functions of spoken discourse markers in the whole range, and the conclusions of each specific function of the six discourse markers under investigation from the previous research provide details that can also match with Brinton's framework. The AntConc presented the concordances for the researcher to analyze the functions of spoken discourse markers in context. Additionally, Relevance Theory was used as an auxiliary theoretical framework in the case of a situation when certain spoken discourse markers' specific function in a certain context was ambiguous.

To answer the fourth question: what are the different pragmatic functions of spoken discourse markers used by Thai EFL learners compared to native English speakers, each spoken discourse marker investigated in this research had comparisons between Thai EFL learners and native English speakers from three facets: a comparison of the raw frequency, its corresponding proportion and the normalized frequency of each function in each spoken discourse markers performed by Thai EFL learners to native English speakers; a comparison of the LL test result of each

118

function in each spoken discourse marker performed by Thai EFL learners to native English speakers; a comparison of the number of participants and its corresponding proportion of the participants who used each function in each spoken discourse marker by Thai EFL learners to native English speakers. The differences were examined comprehensively by the comparisons from three facets.

Chapter 4

Research findings and discussions

This chapter presents the findings in correspondence with the four research questions. By using the quantitative analysis, it shows the findings of the first and the second questions: whether there is a significant difference in use of the spoken discourse markers by Thai EFL learners compared to native English speakers, and whether there is a significant difference in use of the spoken discourse markers between Thai B1-level EFL learners and Thai C1-level EFL learners. By mainly using the qualitative analysis, the research presents the findings of the third, in which the research identifies the functions of each discourse marker used by both Thai B1-level EFL learners and Thai C1-level EFL learners respectively. By using the quantitative analysis, the research presents the main differences in the pragmatic functions of the spoken discourse markers used by Thai EFL learners compared to the functions that have been found in the previous research for the fourth research question. The details of the findings are presented as follows.

4.1 Basic information of learner corpora and SBCSAE

This research built up one learner corpus entitled Thai EFL Learners Spoken English Corpus (TELSEC). It includes two sub-corpora: Thai B1-level EFL learners Spoken English Corpus (BTELSEC) and Thai C1-level EFL learners Spoken English Corpus (CTELSEC). This part mainly presents the basic information of each corpus in this research, including the tokens of each corpus and the number of instances of each spoken discourse marker found in each corpus.

BTELSEC has 25,599 tokens in total. Table 9 summarizes the use of each spoken discourse marker by Thai B1-level EFL learners.

Table 9: Use of each spoken discourse marker in BTELSEC

Spoken discourse markers	As a spoken discourse marker	As a non-spoken discourse marker	Total use of the token	All the other tokens excluded the corresponding spoken discourse marker
like	39	401	440	25,560
so	38	170	208	25,561
well	0	14	14	25,599
you know	22	9	31	25,577
I think	239	6	245	25,360
I mean	8	9	17	25,591
total	346	609	955	25,253

As illustrated in the table above, based on the criteria of identifying the spoken discourse markers in this research, 346 spoken discourse markers were identified in BTELSEC.

It can be seen that Thai B1-level EFL learners used the spoken discourse marker *I think* the most with 239 instances, while no *well* was identified as a discourse marker performed by Thai B1-level EFL learners.

For the use of *like* and *so*, 3 *like* and 2 *so* were identified as the redundant words between two utterances in that the speaker simply repeats them without any semantic or pragmatic sense.

CTELSEC has 49,556 tokens in total. Table 10 summarizes the use of each spoken discourse marker by Thai C1-level EFL learners.

Table 10: Use of each spoken discourse marker in CTELSEC

Spoken	As a spoken	As a non-	Total use of	All the other tokens

discourse markers	discourse marker	spoken discourse marker	the token	excluded the corresponding discourse marker
like	390	966	1356	49,166
so	150	588	738	49,406
well	24	28	52	49,532
you know	68	59	127	49,488
I think	206	295	501	49,350
I mean	116	49	165	49,440
total	954	1985	2939	48,602

As illustrated in the table above, Thai C1-level EFL learners used the spoken discourse marker *like* the most. 30 *like* were identified as the redundant words between two utterances. The speaker repeated *like* several times between two utterances so that these *like* did not own any semantic or pragmatic sense. The use of *I think*, *so* and *I mean* as spoken discourse markers was also comparatively higher than others. 13 *I mean* and 3 *so* were identified the redundant words between two utterances.

Thai C1-level EFL learners performed *well* and *you know* comparatively lower than others. In CTELSEC, there are 52 *well*. 28 *well* were used with their semantic sense in the utterances. 24 *well* were identified as discourse markers based on the given criteria.

127 *you know* were found in CTELSEC. 59 of them were one part of the utterance to serve semantic sense. Thus, they cannot be identified as discourse markers. 68 *you know* were identified as discourse markers finally.

According to the data from the two sub-corpora: BTELSEC and CTELSEC, the research has the raw data of the spoken discourse markers in the learner corpus TELSEC.

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. Table 11 summarizes the use of each spoken discourse marker by Thai EFL learners.

Table 11: Use of each spoken discourse marker in TELSEC

Spoken discourse markers	As a spoken discourse marker	As a non-spoken discourse marker	Total use of the token	All the other tokens excluded the corresponding discourse marker
like	429	1367	1796	74,726
so	188	758	946	74,967
well	24	42	66	75,131
you know	90	68	158	75,065
I think	445	301	746	74,710
I mean	124	58	182	75,031
total	1,300	2594	3894	73,855

Based on the basic information about SBCSAE (Du Bois et al., 2000-2005) and the previous studies that have used SBCSAE (Polat, 2011), SBCSAE has approximately 249,000 tokens. Even though previous research has provided the frequencies of some spoken discourse markers under investigation in this research (Polat, 2011), this research still uses the same criteria mentioned above to identify the spoken discourse markers in SBCSAE or to examine the results provided by the previous research. Table 12 summarizes the use of each spoken discourse marker by native English speakers in SBCSAE.

Table 12: Use of each spoken discourse marker in SBCSAE

Spoken discourse markers	As a spoken discourse marker	As a non-spoken discourse marker	Total use of the token	All the other tokens excluded the corresponding discourse marker
like	1469	800	2269	247,531
so	1743	404	2147	247,257
well	1394	86	1480	247,606
you know	1444	217	1661	247,556
I think	498	23	521	248,502
I mean	548	5	553	248,452
total	7096	1535	8631	241,904

Table 13 presents the raw frequency and the proportion of each spoken discourse marker in each corpus for data analysis.

Table 13: Raw frequency and proportion of each spoken discourse marker in each corpus

Corpora / Discourse markers	Raw frequency and proportion (%) in BTELSEC	Raw frequency and proportion (%) in CTELSEC	Raw frequency and proportion (%) in TELSEC	Raw frequency and proportion (%) in SBCSAE
like	39 (11.3)	390 (40.9)	429 (33.0)	1469 (20.7)
so	38 (11.0)	150 (15.7)	188 (14.5)	1743 (24.6)
well	0 (0.0)	24 (2.5)	24 (1.8)	1394 (19.6)
you know	22 (6.4)	68 (7.1)	90 (6.9)	1444 (20.3)
I think	239 (69.1)	206 (21.6)	445 (34.2)	498 (7.0)
I mean	8 (2.3)	116 (12.2)	124 (9.5)	548 (7.7)
total discourse	346 (100)	954 (100)	1,300 (100)	7096 (100)

markers				

4.2 Question 1: Is there a significant difference in use of spoken discourse markers between Thai EFL learners and native English speakers?

Concerning the complete answer to this question, it involves two comparisons: a comparison of all the spoken discourse markers between each learner corpus and native English speaker corpus, and a comparison of each spoken discourse marker between each learner corpus and native English speaker corpus. The comparisons were done for three pairs of corpora: TELSEC-SBCSAE; BTELSEC-SBCSAE; CTELSEC-SBCSAE.

Based on the methodology, this research used the LL test to have the significant difference. The research adopted the LL critical value in 15.13 at 0.01% level (p<0.0001). This is because according to the previous research (Rayson et al, 2004), the LL critical value of 15.13 is the safest and most recommended critical value when comparing two unbalanced sized corpora.

4.2.1 Comparison in the use of the spoken discourse markers between TELSEC and SBCSAE

This part compares the significant difference in the use of the total six spoken discourse markers and the use of each spoken discourse markers between TELSEC and SBCSAE.

According to the raw data of TELSEC and SBCSAE, the research used the LL test to have the significant difference.

Table 14 illustrates the raw data to do the significance test of the total six spoken discourse markers between TELSEC and SBCSAE.

Table 14: Raw frequency of each spoken discourse marker and non-spoken discourse markers in TELSEC and SBCSAE

	like	so	well	You know	I think	I mean	all non-discourse markers	total
TELSEC	429	188	24	90	445	124	73855	75155
SBCSAE	1469	1743	1394	1444	498	548	241904	249000

Table 15 illustrates the LL test result calculated from Hardie's LL test website.

Table 15: LL test result of the comparison of the frequency of the total six spoken discourse markers between TELSEC and SBCSAE

	Value	df	p-value
LL	1414.14	6	0.0000

Table 16 presents the LL test results in comparison of the use of each spoken discourse marker between TELSEC and SBCSAE retrieved from LL Wizard by Rayson (p<0.0001, df=1).

In the table, O1 and O2 are 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 16: LL test result in 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 |

First of all, as the result shows, there was a significant difference in the use of the total six spoken discourse markers between Thai EFL learners and native English speakers (LL=1414.14, p=0.0000). As is shown in the descriptive data and the result of using LL, it can be concluded that Thai EFL learners underused the spoken discourse markers compared to native English speakers. This result is aligned with the previous research, such as Aijmer (2011, 2016), Polat (2011), Buysse (2012), Diskin (2017), etc., stating the situation where non-native English speakers or EFL learners underused the spoken discourse markers in oral communication.

For the LL test result of each spoken discourse marker, two of them, namely *like* and *I mean*, did not have significant differences. Four spoken discourse markers (*so*, *well*, *you know*, *I think*) had the significant differences between TELSEC and SBCSAE. Three of them were underused (*so*, *well*, *you know*), while the discourse marker *I think* was overused by Thai EFL learners.

The discourse marker *well* had the largest significant difference. After the LL test, it shows that there was a significant difference in the use of the spoken discourse marker *well* between Thai EFL learners and native English speakers (LL=562.17). As is shown in the descriptive data and the LL test result, it can be concluded that Thai EFL learners underused discourse marker *well* compared to native English speakers.

Based on the previous research, *well* has been seen as a focal discourse marker that mainly serves its interpersonal functions in oral communication (Aijmer, 2011, 2016; Brinton, 2008). In line with the findings of Aijmer (2011), it was found that EFL learners underused discourse marker *well*, or they did not even use it as a discourse marker because EFL learners lacked the awareness of using *well* as a discourse marker. Meanwhile, Polat (2011) also discovered the same result,

127

illustrating that non-native English speakers did not use *well* as a discourse marker in conversation.

This finding is also in line with Arya (2020) and Nookam (2010), wherein Thai EFL learners were used in the examination of using the spoken discourse markers. It was found that no spoken discourse marker *well* was used by Thai EFL learners (Nookam, 2010), indicating a deficiency in the use of *well* as a spoken discourse marker by Thai EFL learners.

The discourse marker *you know* also had a relatively large significant difference. The result shows that there was a significant difference in the use of the spoken discourse marker *you know* between Thai EFL learners and native English speakers (LL=339.80). As is shown in the descriptive data and the LL test result, it can be concluded that Thai EFL learners underused the discourse marker *you know* compared to native English speakers. Based on the 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 indicates that Thai EFL learners lack the awareness of using it to have interpersonal sense in English conversation.

The LL test result for the discourse marker *so* shows that there was a significant difference in the use of the spoken discourse marker *so* between Thai EFL learners and native English speakers (LL=236.18). As is shown in the descriptive data and the LL test result, it can be concluded that Thai EFL learners underused the discourse marker *so* compared to native English speakers. This result is aligned with the study of Buysse (2012), which specifically investigated *so* as a discourse marker used by EFL learners. It shows the same result that EFL learners underused the discourse marker *so* in oral communication.

One point worth mentioning here is that based on Arya (2020) and Nookam (2010), it was found that Thai EFL learners used *so* in a relatively high frequency

128

among the spoken discourse markers that were used by Thai EFL learners. However, under the circumstance where a comparison of using *so* was made between Thai EFL learners and native English speakers, the spoken discourse marker *so* was underused by Thai EFL learners.

Another discourse marker that had a significant difference was *I think.* After the LL test, it shows that there was a significant difference in use of the spoken discourse marker *I think* between Thai EFL learners and native English speakers (LL=259.30). As is shown in the descriptive data and the LL test result, it can be concluded that EFL learners overused the discourse marker *I think* compared to native English speakers. The discourse marker *I think* was the only discourse marker under investigation that was overused by Thai EFL learners.

In contrast to the results above, two discourse markers, namely *like* and *I mean*, did not have the significant differences between TELSEC and SBCSAE.

Based on the LL test result, it shows that there was no significant difference in the use of the spoken discourse marker *like* between Thai EFL learners and native English speakers (LL=0.36). The discourse marker *like* was the only single-word discourse marker under investigation in this research that had no significant difference between Thai EFL learners and native English speakers. As the detail of the data will be shown in the following parts, it can be discovered that C1-level learners used *like* in a much higher frequency than B1-level learners.

In the meantime, the LL test result also shows that there was no significant difference in the use of the spoken discourse marker *I mean* between Thai EFL learners and native English speakers (LL=8.90). The discourse marker *I mean* was the only multi-word discourse marker under investigation that had no significant difference between Thai EFL learners and native English speakers. As the data will be

shown in detail in the following parts, C1-level learners used *I mean* as a discourse marker in a much higher frequency than B1-level learners.

In conclusion, the result of the research question 1 illustrated that Thai EFL learners underused the spoken discourse markers in general. Among the six spoken discourse markers under investigation, the discourse markers *well*, *so* and *you know* were underused by Thai EFL learners. The discourse marker *I think* was overused, while the discourse markers *like* and *I mean* did not have the significant differences.

In the following parts, the research presents the comparisons in use of spoken discourse markers between Thai EFL learners with two English proficiency levels and native English speakers. The comparison between Thai B1-level EFL learners and native English speakers is presented first.

4.2.2 Comparison in the use of the spoken discourse markers between BTELSEC and SBCSAE

This part compares the significant difference in the use of the total six spoken discourse markers and the use of each spoken discourse markers between BTELSEC and SBCSAE.

Table 17 illustrates the raw data to do the significance test of the total six spoken discourse markers between BTELSEC and SBCSAE.

Table 17: Raw frequency of each spoken discourse marker and non-spoken discourse markers in BTELSEC and SBCSAE

	like	*so*	*well*	*You know*	*I think*	*I mean*	all non-discourse markers	total
BTELSEC	39	38	0	22	239	8	25253	25599
SBCSAE	1469	1743	1394	1444	498	548	241904	249000

130

Table 18 illustrates the LL test result calculated from Hardie's LL test website.

Table 18: LL test result of the comparison of the frequency of the total six spoken

discourse markers between BTELSEC and SBCSAE

	Value	**df**	**p-value**
LL	1065.56	6	0.0000

Table 19 presents the LL test result in comparison of the use of each spoken discourse markers between BTELSEC and SBCSAE.

Table 19: LL test result in comparison of the use of each spoken discourse marker

between BTELSEC and SBCSAE

	O1	**%1**	**O2**	**%2**		**LL**
like	39	0.15	1469	0.59	-	110.51
so	38	0.15	1743	0.70	-	153.88
well	0	0.00	1394	0.56	-	272.83
you know	22	0.09	1444	0.58	-	158.58
I think	239	0.93	498	0.20	+	302.94
I mean	8	0.03	548	0.22	-	61.47

First of all, as the result shows, there was a significant difference in the use of the total six spoken discourse markers between Thai B1-level EFL learners and native English speakers (LL=1065.56, p=0.0000). As is shown in the descriptive data and the LL test result, it can be concluded that Thai B1-level EFL learners underused discourse markers compared to native English speakers.

As is shown in both the descriptive data and the inferential data, Thai B1-level EFL learners performed the spoken discourse markers much less frequently than native English speakers. It indicates that Thai B1-level EFL learners were not so

familiar with the use of spoken discourse markers in English conversations. This finding is aligned with several previous studies, such as Aijmer (2004, 2011), Polat (2011), Diskin (2017), etc., reflecting the similar problem that lower-level EFL learners used spoken discourse markers at a low frequency.

Besides the comparison of the total six spoken discourse markers, each discourse marker had a significant difference between BTELSEC and SBCSAE. As demonstrated by the table above, Thai B1-level EFL learners underused five spoken discourse markers and overused one spoken discourse marker.

Similar to the situation of TELSEC, the discourse marker *well* in BTELSEC also had a large significant difference compared to SBCSAE. The result shows that there was no use of the discourse marker *well* in BTELSEC. Hence, there was a significant difference in the use of the spoken discourse marker *well* between Thai B1-level EFL learners and native English speakers (LL=272.83). As is shown in the descriptive data and the LL test result, it can be seen that Thai B1-level EFL learners lacked the awareness of using discourse marker *well* in English conversation.

The situation of discourse marker *well* is unique compared to other discourse markers under investigation in this research. The significant difference is quite huge compared to the other two single-word discourse markers under investigation in this research. This situation does not exist alone, which can be found in the previous research. In Polat (2011)'s research, the participant in the case study did not produce any *well* as a discourse marker as well. It showed the phenomenon that the participant did not have the relative knowledge of how to use *well* as a discourse marker so that the participant never had the awareness of using *well* as a discourse marker in English conversation. Previous research (Aijmer, 2011) also found that non-native English speakers used *well* as a discourse marker in a low frequency. It is thus conceivable to have this result in this research.

The discourse marker *you know* also had a relatively large significant difference. The result shows that there was a significant difference in use of the spoken discourse marker *you know* between Thai B1-level EFL learners and native English speakers (LL=158.58). As is shown in the descriptive data and the LL test result, it can be concluded that Thai B1-level EFL learners had a tendency that they underused the discourse marker *you know* compared to native English speakers.

Although the significant difference is huge compared to native English speakers, it can be seen that Thai B1-level EFL learners used *you know* as a discourse marker in English conversation. It should be noted that all the production of *you know* comes from the same B1-level participant in BTELSEC. This result is aligned with the findings of Hellermann and Vergun (2007), which also concluded that lower-level EFL learners did not quite often use *you know* as a discourse marker. It is assumed that lower-level EFL learners had not acquired the knowledge of how to use discourse markers in English conversation, and their English proficiency level had not been high enough for them to produce discourse markers.

The LL test results for the discourse markers *so* and *like* are 153.88 and 110.51, respectively. It illustrates that Thai B1-level EFL learners underused both spoken discourse markers.

For the situation of the discourse marker *so*, it can be seen that Thai B1-level EFL learners still used the discourse marker *so* in English conversation, yet not as frequently as native English speakers. It is aligned with the findings of Buysse (2012), discovering that EFL learners used *so* as a discourse marker at a much lower frequency than native English speakers.

The same as the situation of *so*, it was found that Thai B1-level EFL learners used the discourse marker *like*, but not as often as native English speakers. It is

aligned with the findings of Diskin (2017), which illustrated that EFL learners used *like* as a discourse marker at a much lower frequency than native English speakers.

Regarding the use of the discourse marker *I mean*, the result of LL is 61.47. It shows that Thai B1-level EFL learners underused the discourse marker *I mean* compared to native English speakers. The significant difference of *I mean* is the smallest one among the six discourse markers under investigation in this research. According to the descriptive data, Thai B1-level EFL learners still used *I mean* as a discourse marker in English conversation, but the frequency was very low. When looking at the details of the data, only two participants in B1-level used *I mean* as a discourse marker in English conversation. It indicates that Thai B1-level EFL learners were lack knowledge or awareness of using *I mean* as a discourse marker in English conversation.

In contrast to the five discourse markers above, the situation of the use of the discourse marker *I think* by Thai B1-level EFL learners is different. The result shows that there is a significant difference in the use of the spoken discourse marker *I think* between Thai B1-level EFL learners and native English speakers (LL=302.94). As is shown in the descriptive data and the result of using LL, it can be concluded that Thai B1-level EFL learners had a tendency that they overused the discourse marker *I think* compared to native English speakers.

As is shown in the descriptive data, native English speakers used the discourse marker *I think* much less than Thai B1-level EFL learners. The significant difference is relatively huge compared to the situation of other discourse markers in this research. Among the three multi-word discourse markers in this research, the significant difference of discourse marker *I think* is the largest one. It indicates the fact that Thai B1-level EFL learners tended to rely on using *I think* too much.

In conclusion, Thai B1-level EFL learners underused the spoken discourse markers compared to native English speakers in general. According to the descriptive data and the LL test result, Thai B1-level EFL learners tended to underuse five spoken discourse markers under investigation. Moreover, no case of using the spoken discourse marker *well* by Thai B1-level EFL learners was found. Besides, they overused the discourse marker *I think*.

In the next part, the comparison between Thai C1-level EFL learners and native English speakers is presented.

4.2.3 Comparison in the use of the spoken discourse markers between CTELSEC and SBCSAE

This part compares the significant difference in the use of the total six spoken discourse markers and the use of each spoken discourse markers between CTELSEC and SBCSAE.

Table 20 illustrates the raw data to do the significance test of the total six spoken discourse markers between CTELSEC and SBCSAE.

Table 20: Raw frequency of each spoken discourse marker and non-spoken discourse markers in CTELSEC and SBCSAE

	like	*so*	*well*	*You know*	*I think*	*I mean*	all non-discourse markers	total
CTELSEC	390	150	24	68	206	116	48602	49556
SBCSAE	1469	1743	1394	1444	498	548	241904	249000

Table 21 illustrates the LL test result calculated from Hardie's LL test website.

Table 21: LL test result of the comparison of the frequency of the total six spoken discourse markers between CTELSEC and SBCSAE

	Value	df	p-value
LL	783.39	6	0.0000

Table 22 presents the LL test result in comparison of the use of the spoken discourse markers between CTELSEC and SBCSAE.

Table 22: LL test result in comparison of the use of each spoken discourse marker between CTELSEC and SBCSAE

	O1	%1	O2	%2		LL
like	390	0.79	1469	0.59	+	24.17
so	150	0.30	1743	0.70	-	123.11
well	24	0.05	1394	0.56	-	348.86
you know	68	0.14	1444	0.58	-	213.70
I think	206	0.42	498	0.20	+	69.57
I mean	116	0.23	548	0.22	+	0.36

As the LL test result shows, there was a significant difference in the use of the total six spoken discourse markers between Thai C1-level EFL learners and native English speakers (LL=783.39, p=0.0000). As is shown in the descriptive data and the LL test result, it can be concluded that Thai C1-level EFL learners underused discourse markers compared to native English speakers even though the significant difference was smaller than Thai B1-level EFL learners.

For the situation of each discourse marker used by Thai C1-level EFL learners, they underused three discourse markers: *well*, *you know* and *so*. In the meantime, they overused two discourse markers *I think* and *like*. There was no significant difference in the use of the discourse marker *I mean*.

The situation in the use of the discourse marker *well* by Thai C1-level EFL learners was still the worst one. The result of LL shows that there was a significant difference in the use of the spoken discourse marker *well* between Thai C1-level EFL learners and native English speakers (LL=348.86). As is shown in the descriptive data and the LL test result, it can be concluded that Thai C1-level EFL learners underused the discourse marker *well* compared to native English speakers.

For the situation of the discourse marker *you know* used by Thai C1-level EFL learners, the LL test result shows that there was a significant difference in use of the spoken discourse marker *you know* between Thai C1-level EFL learners and native English speakers (LL=213.70). As is shown in the descriptive data and the LL test result, it can be concluded that Thai C1-level EFL learners underused the discourse marker *you know* compared to native English speakers. The significant difference was relatively huge compared to the situation of other discourse markers in this research.

The LL test result for the use of the discourse marker *so* also shows that there was a significant difference in the use of the spoken discourse marker *so* between Thai C1-level EFL learners and native English speakers (LL=123.11). As is shown in the descriptive data and the LL test result, it can be concluded that Thai C1-level EFL learners underused the discourse marker *so* compared to native English speakers.

In contrast to the results above, for the situation in the use of the discourse marker *I think* in CTELSEC, the result shows that there was a significant difference in the use of the spoken discourse marker *I think* between Thai C1-level EFL learners and native English speakers (LL=69.57). As is shown in the descriptive data and the LL test result, it can be concluded that Thai C1-level EFL learners overused the discourse marker *I think* than native English speakers.

As similar as the situation of the discourse marker *I think*, the LL test result for the discourse marker *like* used by Thai C1-level EFL learners also shows that there

137

was a significant difference in the use of the spoken discourse marker *like* between Thai C1-level EFL learners and native English speakers (LL=24.17). As is shown in the descriptive data and the LL test result, it can be concluded that Thai C1-level EFL learners overused the discourse marker *like* compared to native English speakers.

The use of the discourse marker *I mean* by Thai C1-level learners was almost the same as native English speakers based on the result of LL (LL=0.36). There was no significant difference in the use of *I mean* between the two groups.

In conclusion, Thai C1-level EFL learners underused three spoken discourse markers: *well*, *you know* and *so*. Moreover, they overused two spoken discourse markers: *I think* and *like*. There was no significant difference in using the discourse marker *I mean*.

The previous parts have done the comparisons in the use of spoken discourse markers between Thai EFL learners with B1-level and C1-level and native English speakers. In the next part, the significant difference in use of the spoken discourse markers between Thai B1-level and Thai C1-level EFL learners will be illustrated.

4.3 Question 2: Is there a significant difference in use of spoken discourse markers between Thai B1-level EFL learners and Thai C1-level EFL learners?

To answer this question, two steps are processed: a significance test in the use of the total six discourse markers between BTELSEC and CTELSEC, and a significance test in the use of each spoken discourse marker between both corpora.

Table 23 illustrates the raw data to do the significance test of the total six spoken discourse markers between TELSEC and SBCSAE.

Table 23: Raw frequency of each spoken discourse marker and non-spoken discourse markers in BTELSEC and CTELSEC

	like	so	well	You know	I think	I mean	all non-discourse markers	total
BTELSEC	39	38	0	22	239	8	25253	25599
CTELSEC	429	188	24	90	445	124	73855	75155

Table 24 illustrates the LL test result calculated from Hardie's LL test website.

Table 24: LL test result of the comparison of the frequency of the total six spoken discourse markers between BTELSEC and CTELSEC

	Value	df	p-value
LL	315.92	6	0.0000

Table 25 presents the LL test result in comparison of the use of the spoken discourse markers between BTELSEC and CTELSEC.

Table 25: LL test result in comparison of the use of each spoken discourse marker between BTELSEC and CTELSEC

	O1	%1	O2	%2		LL
like	39	0.15	390	0.79	-	147.46
so	38	0.15	150	0.30	-	17.53
well	0	0.00	24	0.05	-	19.99
you know	22	0.09	68	0.14	-	3.92
I think	239	0.93	206	0.42	+	71.93
I mean	8	0.03	116	0.23	-	54.52

Based on the raw data, Thai B1-level EFL learners produced 346 discourse markers under investigation in this research, while Thai C1-level EFL learners

produced 954 discourse markers. The LL test result shows that there was a significant difference in the use of the total six spoken discourse markers between Thai B1-level EFL learners and Thai C1-level EFL learners (LL=315.92, P=0.0000).

Based on the LL test result, it can be seen that there was a significant improvement in using the spoken discourse markers with the improvement of the English level from B1-level to C1-level. Thai C1-level EFL learners used the spoken discourse markers much more often than Thai B1-level learners. However, due to the LL test result in the first question that both Thai B1-level and C1-level learners underused the spoken discourse markers compared to native English speakers, the improvement of using the spoken discourse markers still had not reached the level of how native English speakers used the spoken discourse markers in English conversation.

As the table above illustrated, the use of the spoken discourse marker *you know* between the two levels of Thai EFL learners had no significant difference (LL=3.92). It indicates that with the improvement of the English proficiency from B1-level to C1-level, Thai EFL learners had little progress in performing *you know* as a spoken discourse marker in English conversation.

In contrast, the other five spoken discourse markers had the significant differences, wherein the use of the spoken discourse marker *like* had the largest LL test result. Compared to Thai C1-level EFL learners, the only spoken discourse marker that was overused by Thai B1-level EFL learners was *I think*, whereas the other four spoken discourse markers were underused by Thai B1-level EFL learners.

Based on the LL test result, it is evident to see that Thai C1-level EFL learners used the discourse marker *like* much more often than Thai B1-level EFL learners (LL=147.46). As is shown in both the inferential data and the descriptive data in the analysis of the first question, Thai C1-level EFL learners overused the discourse

markers *like*. It is therefore interesting to see that with the improvement of the English proficiency from B1-level to C1-level, Thai EFL learners performed the discourse marker *like* from infrequent use to overuse it in conversation.

With regards to the use of *I think*, it can be seen that Thai B1-level EFL learners used the discourse marker *I think* more often than Thai C1-level EFL learners. The result indicates that with the increase in the English level, Thai EFL learners tended to insert fewer *I think* as a discourse marker in English conversation. However, based on the result of the first question, both levels still used the discourse marker *I think* too much compared to native English speakers.

Similar to the situation of the discourse marker *like*, the use of the discourse marker *I mean* also showed a significant difference between the two levels (LL=54.52). The result of question 1 showed that Thai B1-level EFL learners underused *I mean*, while Thai C1-level EFL learners used it at a similar frequency as native English speakers. Hence, combining the findings of question 1 and question 2, it was found that with the development of the English proficiency from B1-level to C1-level, Thai EFL learners had a significant improvement in using the discourse marker *I mean* frequently similar to native English speakers.

As for the use of *so* and *well*, it shows that there was a significant difference in the use of both spoken discourse markers between B1-level learners and C1-level learners (LL=17.53 and 19.99, respectively). For the situation of using the spoken discourse marker *so*, Thai EFL learners tended to use it more often with the development of their English level. For the use of the spoken discourse marker *well*, it was as unique as had been mentioned above. For one thing, Thai B1-level EFL learners did not produce any of it found in this research. For another, Thai C1-level EFL learners rarely produced *well* as a spoken discourse marker in English conversation. Thus, both levels of Thai EFL learners seem underused *well* as a

discourse marker with a seriously large gap compared to native English speakers. However, it is encouraging to see that there was a significant improvement in using *well* from B1-level to C1-level.

It is noticed from the data that the trend of using the spoken discourse markers was towards a better direction comprehensively from a lower level to a higher level in English proficiency based on the raw frequencies, the normalized frequencies and the LL test results of each spoke discourse marker and the total six spoken discourse markers. Besides, the number of Thai C1-level participants who used each spoken discourse marker also basically increased. Table 26 presents the number of participants who use each spoken discourse marker in this research.

Table 26: Number of participants who use each spoken discourse marker

Discourse marker	Number of Thai B1-level EFL learners who use the corresponding spoken discourse marker	Number of Thai C1-level EFL learners who use the corresponding spoken discourse marker
like	12	16
so	10	14
well	0	3
you know	1	4
I think	30/9	30/2
I mean	2	13
Total	30/14	30/20

It must be noted here as two numbers are shown in the table above in the columns *I think* and the total spoken discourse markers. This is because all the Thai participants in this research used the function of *I think*, namely, marking opinion, feeling or stance. Hence, it may not be objective and comprehensive to only show the

number "30" in each column. Under the circumstances, this research also presents the number of participants who used the functions excluded the function of *I think*, marking opinion, feeling or stance, in order to show the results more objectively and comprehensively.

As the table shows above, except for *I think*, the number of Thai C1-level EFL learners used each spoken discourse marker more than Thai B1-level EFL learners, making a total of 30% increase in using all the spoken discourse markers in this research.

Overall, it is observed that with the improvement of the English proficiency, the performance of the spoken discourse markers increased proportionally, which is in line with the findings of several previous research (Asik & Cephe, 2013; Hellermann & Vergun, 2007; Lee-Goldman, 2010). It was illustrated, as cited by Hellermann and Vergun (2007), that language learners with a higher proficiency were more likely to perform spoken discourse markers. Supported by Asik and Cephe (2013), it was found that higher proficiency level or upper-level students used more focal discourse markers in conversation. From another perspective, it was also argued that using spoken discourse markers in conversation may influence non-native English speakers' fluency (Asik & Cephe, 2013). For one thing, it was found discourse markers, as their multi-faceted functions in different contexts, were used as a lubricant in interaction to reduce the potential difficulty in understanding (Yang, 2011). For another, it was argued that using a proper right amount of spoken discourse markers in communication may also overcome the overuse of fillers, such as *uhh*, *emm*, resulting in a more fluent conversational interaction (Asik & Cephe, 2013). From the two perspectives pointed out above, it can be assumed that there is a relationship between English proficiency and performance of the spoken discourse markers in communication, wherein a higher level of English proficiency can bring out more

143

discourse markers, while the use of some spoken discourse markers, on the contrary, has an effective impact on the fluency in interaction.

Besides the implication illustrated above, three points are worth mentioning here concerning the findings of the comparison in the performance of the spoken discourse markers between the two levels of Thai EFL learners.

It must be pointed out, in the first place, that the performance of the spoken discourse markers by individuals involved in this research varied largely and unevenly. Regarding the situation of B1-level participants in this research, for instance, all the production of *you know* came from the same participant; the performance of *I mean* only came from two participants. As for the situation of C1-level participants in this research, only three participants used the spoken discourse marker *well*; four participants used the discourse marker *you know*; thirteen Thai C1-level participants, which were fewer than half of the Thai C1-level participants, used the discourse marker *I mean*. In the meantime, it was also observed that 9 participants tended to use some discourse markers repeatedly through the dialogic interaction or only focused on using one spoken discourse marker from the beginning until the end of the conversation. This phenomenon does not stand alone, but has been found in previous studies as well (Aijmer, 2011; Polat, 2011). Aijmer (2011) illustrated that large differences in using the discourse markers were found when the data was broken down to each learner, suggesting that different learners acquired the discourse markers in different ways and to different extents. It was also pointed out by Polat (2011) that different language learners tended to only acquire the discourse markers that were more readily to them, resulting in an uneven distribution of the performance of the discourse markers. In line with the findings of these previous studies, it is assumed that Thai EFL learners acquired the spoken discourse markers in an unsystematic way, leading to their large differences and uneven use of them. Some

learners picked up the ones that were exposed to them in a natural linguistic environment, while others never realized this category because of the deficiency of being exposed to them.

The point stated above brings to the second concern that even though there was a significant improvement in the performance in five spoken discourse markers investigated in this research, only one of them, namely *I mean*, reached a similar frequency as native English speakers, let alone the condition where there was no significant improvement of the performance of *you know* from B1-level to C1-level. It reveals that Thai EFL learners may face a dearth of linguistic input from the pragmatic point of view so that the discrepancy still exists when they are at high English level, with inadequacy in the ability to use English from cultural, social and situational perspectives (Fung & Carter, 2007).

Another interesting phenomenon worth discussing here is the performance of *like* from underuse by Thai B1-level EFL learners to overuse by Thai C1-level EFL learners, reflecting a procedure from a low performance to an over-reliance on one discourse marker. As studied by Diskin (2017), it was discovered that non-native English speakers may achieve the use of the discourse marker *like* similar to native English speakers in its frequency driven by the natural exposure to the English environment. Unlike the interpersonal-centered spoken discourse markers such as *well* and *you know*, *like* mainly serve its pragmatic role in communication by its versatile textual functions. Hence, it is assumed that Thai EFL learners are more readily and preferable to acquire *like* as it is easier and more generic to use in communication. It may bring out an implication, combining the rare use of *well* and *you know*, that it is more suitable for EFL learners to start the acquisition of the discourse markers centered by the textual functions. On the contrary, over-reliance should be paid attention to when learners' English proficiency is improved.

145

In conclusion, it was found that there was a significant difference in the use of the spoken discourse markers between Thai B1-level EFL learners and Thai C1-level EFL learners. Thai C1-level EFL learners used the spoken discourse markers more often than B1-level learners. Among the six spoken discourse markers under investigation in this research, Thai B1-level EFL learners used four of them in a much lower frequency, namely *like*, *so*, *well* and *I mean*. However, B1-level learners tended to use the discourse marker *I think* much more often than C1-level learners. In contrast, there was no significant difference of using the discourse marker *you know*.

4.4 Question 3: What are the pragmatic functions of spoken discourse markers used by Thai EFL learners?

Concerning the complete answer to this question, this research presents the pragmatic functions of each spoken discourse marker used by both Thai B1-level EFL learners and Thai C1-level EFL learners, respectively.

Two points must be illustrated before the presentation of the detailed function analysis. The analysis of the pragmatic functions in this research was all completed by the researcher alone based on Brinton's (2008) framework, the conclusions of the pragmatic functions by previous research and Relevance Theory. However, there may exist the subjectivity issue in the analysis of some instances because the function analysis was based on the interpretation from the researcher's perspective. Meanwhile, some previous research (Fraser, 1999) pointed out that in some contexts, there may be more than one pragmatic function of a discourse marker. However, some studies (Schiffrin, 1987) illustrated that a discourse marker had a main function in a specific context. Following this idea, recent studies that analyzed the functions of a spoken discourse marker only presented the main function of a spoken discourse marker in a specific context (Aijmer, 2011; Brinton, 2010; Diskin, 2017). Under the

circumstances, this research also presents the interpretation of the main pragmatic function of a spoken discourse marker in a specific context. However, more analysis or explanation is also presented after the interpretations of some functions where there existed ambiguity.

This research firstly presents the pragmatic functions of each spoken discourse marker used by Thai B1-level EFL learners.

4.4.1 Pragmatic functions of each spoken discourse marker used by Thai B1-level EFL learners

4.4.1.1 Pragmatic functions of *like* as a discourse marker

It is observed that Thai B1-level EFL learners performed six pragmatic functions of the discourse marker *like*. According to Brinton's (2008) classification of the functions of the discourse markers, they all belonged to the textual functions. Table 27 shows the textual functions that were discovered in BTELSEC.

Table 27: Textual function of like as a discourse marker used by Thai B1-level EFL learners

Textual function of *like* as a discourse marker	Number of cases
To exemplify	11
Marking a focus on new information	8
To have a quotation	7
More information/explanation	7
Searching for the following words	5
As an approximator	1

As can be seen, out of 39 instances, 11 of them served to exemplify more details on the previous utterance, resulting in the most widely used function of the discourse marker *like* by Thai B1-level EFL learners. Marking a focus on new information was

the second most-used textual function, followed by the other two functions: to have a quotation and to provide more information or explanation for a previous utterance. Five cases accounted for searching for the following words. Only one case was discovered to use *like* as an approximator.

This research presents how it identified each function of the discourse marker *like* used by Thai B1-level EFL learners in detail.

To exemplify

It was illustrated by Brinton (2008) that the discourse markers were derived from full lexemes with their semantic content through the diachronic study. Hence, it can be interpreted that certain functions of a discourse marker retain traces of its original propositional meaning (Hansen, 1998; Traugott & Dasser, 2002). Supported by the previous research (Müller, 2005; Diskin, 2017), this function is derived from the original semantic meaning of the lexicon *like*, denoting a subordinate (an example) towards a superordinate when *like* is a preposition. Distinguished from this semantic meaning, the exemplifications do not consist with the prior utterance as a direct relationship between subordinate and superordinate upon the examination of the context. The speaker performs this function when he or she provides the exemplifications with details towards a preceding utterance. The speaker uses this function in favor of the hearer in that he or she wants to make sure that the preceding utterance can be fully understood by the hearer. It must be noted that the exemplifications preceded by *like* comprise a series of utterances, wherein more than one piece of information is given. The following two excerpts illustrate this function of the discourse marker *like*. In all excerpts as follows, INT stands for the interlocutor (the researcher) who did the dialogue with each participant in the conversation.

Mr. B1-14: but I still learn math .

Mr. B1-14: (.) and (.) I good at math .

148

INT: ok .

Mr. B1-14: (.) but sometime (.) I feel I have question .

Mr. B1-14: like [DM] (.) why I have to do this .

Mr. B1-14: why I have to do that .

(Mr. B1-14)

Excerpt 1

INT: you mean her facial expression is not so good ?

Ms. B1-17: uhh (…) yeah .

Ms. B1-17: <laugh>

Ms. B1-17: like [DM] (.) she look you (.) not happy .

Ms. B1-17: very very angry .

Ms. B1-17: and (.) maybe you think .

Ms. B1-17: oh .

Ms. B1-17: (.) why [/] why she like that ?

(Ms. B1-17)

Excerpt 2

In except 1, the participant discusses his thought that he has a lot of questions about his life in mind even though he has mastered well on the main subject (math). He thereby states the utterance "I feel I have question" before inserting the discourse marker *like*. To make sure that the hearer can fully understand him, he specifies two concrete questions to present his doubt towards life. The two concrete questions overtly are the exemplifications of the utterance "I feel I have question". By illustrating the exemplifications (the two questions), the speaker denotes his doubt towards life more explicitly.

In excerpt 2, the participant expresses that her best friend does not have a good facial expression in the preceding context. To fulfill the doubt from the hearer, she continues with a series of utterances full of concrete images accounting for the unattractive facial expression of her friend after inserting the discourse marker *like*, i.e., look you not happy, very very angry, etc. It is hence construed that her friend's facial expression is not good in other people's eyes. These concrete images become the exemplifications for the interpretation of the prior utterance "her facial expression is not so good".

It is noted that the exemplifications in all the instances found in BTELSEC comprise several different pieces of information. They all account for exemplifying a prior utterance before the discourse marker *like*. In the meantime, the utterances preceded by *like* cannot be regarded as the direct subordinate to the prior utterance before *like*.

Marking a focus on new information

Illustrated by the previous research (Diskin, 2017; Polat, 2011), the speaker tends to insert *like* before a piece of new information for the purpose of emphasizing it so that the hearer can pay attention to it. A point worthy of note concerning identifying this function is that the emphasized information preceded by *like* must be a new one in a way in which it never appears in the previous utterances. Marked by the discourse marker *like*, the hearer is expected to get the signal and focus on this piece of new information, as illustrated by the two excerpts below.

Ms. B1-09: yeah (.) and I help her .

Ms. B1-09: I remember .

Ms. B1-09: like [DM] (.) after school and weekend

(Ms. B1-09)

Excerpt 3

150

Ms. B1-21: but some teacher .

Ms. B1-21: they &-uhh (.) ask [/] ask student .

Ms. B1-21: write like [DM] diary .

(Ms. B1-21)

Excerpt 4

In both cases, both participants insert the discourse marker *like* before a piece of new information that never appears in the preceding context, respectively. In excerpt 3, the participant tells the hearer that she helped her mom with some work when she did not need to study. The new information "after school and weekend" is added by the speaker by using the discourse marker *like* to emphasize the time slot. In excerpt 4, the participant wants the hearer to focus on the word *diary*, because it is a new experience for her to be asked to keep a diary by her teacher in the university. Hence, the speaker inserts the discourse marker *like* before the word *diary* to signal to the hearer that she thinks that it is special for a teacher to ask students to keep a diary at the university.

As is discovered in both excerpts, the semantic meaning of the new information preceded by *like* and the semantic meaning of the prior utterance before *like* are consistent. Without the discourse marker *like*, there is no marking the emphasis on the new information. It is thereby construed that the discourse marker *like* is deliberately inserted by the speaker to be the usage of emphasis.

To have a quotation

This function is used when learners directly quote certain original utterance. It is relatively simple to identify this function in that there is a quotation directly preceded

by the discourse marker *like*. The excerpt taken from BTELSEC below illustrates this function.

> Ms. B1-09: I know .
>
> Ms. B1-09: it's tired .
>
> Ms. B1-09: yes .
>
> Ms. B1-09: and (.) I always (.) &-emm (..) talk her .
>
> Ms. B1-09: like [DM] why you live so far ?
>
> (Ms. B1-09)

<center>**Excerpt 5**</center>

In excerpt 5, the participant simply quotes the utterance that she asked her friend before, which is preceded by the discourse marker *like*. It is discovered that the frequency of this function is in the middle among all the functions used by Thai B1-level EFL learners, reflecting that Thai B1-level EFL learners have an awareness of using this function in communication.

To provide more information/explanation

It was stated by the previous research (Diskin, 2017; Müller, 2005) that the speaker supplemented the preceding utterance with more information or explanation in favor of a better mutual understanding. The vague semantic meaning in the preceding context in conversation was thereby rendered by the provided more information or explanation through the signal of using the discourse marker *like* (Diskin, 2017; Tagliamonte, 2005). It is worthwhile to note that the provided more information or explanation is regarded as a supplementary material to the preceding utterance, which should be short and simple, i.e., one word, a short expression or a short sentence. The excerpts extracted from BTELSEC illustrate this point.

> Mr. B1-05: sometime I buy something .

<center>152</center>

Mr. B1-05: but (.) not expensive .

Mr. B1-05: I buy in market .

Mr. B1-05: like [DM] night market .

(Mr. B1-05)

Excerpt 6

Mr. B1-18: &-emm (..) sometime I [/] I [/] I just (…) .

Mr. B1-18: <I just> [//] I just feel (.) I [/] I +… (.) .

Mr. B1-18: +/ sometime I also talk to (.) my friend .

Mr. B1-18: (.) like [DM] friend I just tell you .

INT: oh yes .

(Mr. B1-18)

Excerpt 7

In excerpt 6, the participant supplements the information "night" to the prior information "market" to explain that he usually goes shopping in the night market because the stuff in a night market is cheap. It can be interpreted that the speaker consciously wants to reach a better mutual understanding by adding a little more information to the prior utterance.

In excerpt 7, the participant tries to share some information with the hearer about what he would do with his friends. He uses *like* to lead a piece of additional explanation that he likes to chat with his friend, the one who has been discussed in the preceding context.

It should be noted that the provided more information or explanation in both excerpts above is concise, with only one word or a short expression.

Searching for the following words

153

This function is used when the speaker does not know what words he or she is going to say in the following context of conversation. It is also relatively easy to recognize it because the speaker tends to have a hesitation when inserting the discourse marker *like*. Furthermore, the discourse marker *like* tends to co-occur with fillers, i.e., *uhh*, *emm*, *uhhm*, etc. This frequent co-occurrence suggests the hesitation of the speaker. The two excerpts below reveal this point.

Mr. B1-14: sometime &-uhh (..) .

Mr. B1-14: I confuse .

Mr. B1-14: because (.) teacher (.) &-uhh (..) .

Mr. B1-14: like [DM] &-uhh (.) teacher ask question .

Mr. B1-14: but I don't (.) understand

Mr. B1-14: but I want to (.) answer

(Mr. B1-14)

Excerpt 8

Ms. B1-22: I want to study (..) continue .

Ms. B1-22: (..) like [DM] &-uhhm (.) what's that in English ?

INT: Master Degree ?

Ms. B1-22: Master Degree .

Ms. B1-22: yes .

(Ms. B1-22)

Excerpt 9

In excerpt 8, the hesitation of the participant appears after he spells out the word *teacher*, followed by the filler *uhh* and a longer pause than usual. He then inserts *like* and another *uhh* before saying the word *teacher* again. This series of phenomena

reflects the fact that the speaker is thinking the words for the following utterance. In excerpt 9, apart from the co-occurrence of the discourse marker *like* and the filler *uhhm*, the participant directly spells out *what's that in English*, denoting that she was trying to search for the expression *master degree*. The same situation happens twice in BTELSEC.

It is interesting to point out here that EFL learners use this function shared with different discourse markers (i.e. *like*, *I mean*) due to their limited vocabulary or the lack of knowledge of English (Aijmer, 2011). It is hence considered a function developed by non-native English speakers.

As an approximator

Only one case is found in BTELSEC to use *like* as an approximator. The case is as follows.

> INT: when did you go there ?
>
> Ms. B1-26: &-emm (..) about (.) two years ago .
>
> Ms. B1-26: I go with family .
>
> Ms. B1-26: we stay like [DM] three day .
>
> (Ms. B1-26)

Excerpt 10

This function is named after "approximator" because *like* performs as similar as the word "approximate", denoting the meaning of not completely accurate or exact (Diskin, 2017; Müller, 2005). In this condition, it is observed by the previous researchers that the discourse marker *like* co-occured with a number in order to adopt this function (Diskin, 2017; Tagliamonte, 2005). This co-occurrence has been seen as a pattern, supported and acknowledged by different researchers (Aijmer, 2011; Diskin,

2017; Müller, 2005; Tagliamonte, 2005), that the speaker signals the hearer to interpret the number preceded by *like* as a not completely accurate one.

As shown in excerpt 10, the participant is not sure the exact number of the days. Hence, she inserts *like* before the number "three" to show that it is approximately three days rather than exact three days. The low performance of this function reveals that Thai B1-level EFL learners may not be so familiar with this function.

In a short summary, Thai B1-level EFL learners use *like* as a discourse marker mainly focusing on its four textual functions. Meanwhile, they seldom use it as an approximator. No use of its interpersonal function is discovered.

4.4.1.2 Pragmatic functions of *so* as a discourse marker

Unlike the situation of the discourse marker *like*, Thai B1-level EFL learners used *so* as a spoken discourse marker for both textual functions and interpersonal functions. Based on the original data, it is discovered that Thai B1-level EFL learners used five textual functions and three interpersonal functions. Table 28 and Table 29 summarize the functions that Thai B1-level EFL learners used in English conversation.

Table 28: Textual functions of so as a discourse marker used by Thai B1-level EFL learners

Textual functions of *so* as a discourse marker	Number of cases in the raw data
Marking a result/consequence	11
A transition sequentially	8
Marking a start of a narration	5
Topic shift back	3
Leading a summary	2

Table 29: Interpersonal functions of so as a discourse marker used by Thai B1-level EFL learners

Interpersonal functions of *so* as a discourse marker	Number of cases in the raw data
Marking opinion/feeling/stance	7
Prompting the hearer to speak/an implied result	1
hold the floor	1

Both tables above illustrate the phenomenon that the textual functions of the discourse marker *so* were used more often than its interpersonal functions. For its textual functions, marking a result or consequence was performed the most by B1 learners. Its function as a transition sequentially in conversation was also performed at a high frequency. Few instances were identified in the other three textual functions. For its interpersonal functions, it is notable that marking opinions, feelings or stance owns the predominant interpersonal function, whereas the other two types of interpersonal functions only have one case, respectively.

This research presents the details of each function of *so* as a discourse marker used by Thai B1-level EFL learners as follows.

Marking a result/consequence

The same as the case of the discourse marker *like*, the discourse marker *so* also has its pragmatic function that originated from its semantic meaning. It is believed that this function, namely marking a result or a consequence, is derived from the original semantic meaning of the lexicon *so*. (Blakemore, 2014; Brinton, 1996, 2008; Buysse, 2012). As is known, the main semantic meaning of the lexicon *so* is to lead a statement as a direct result or consequence of some cause. It was illustrated that this semantic meaning should yield a direct relationship between cause and effect (Müller, 2005). In other words, this semantic meaning of the lexicon *so* conforms to the contextual environment where a direct cause-and-effect logic exists. In contrast, another contextual environment involving *so* was found that there was no direct

cause-and-effect relationship, wherein the statement before *so* did not lead to the result or consequence after *so*, resulting in an unnecessary but sufficient condition in context. As such, it is conceivable that *so* only promoted a textual function referring to marking a result or a consequence (Buysse, 2012; Müller, 2005). The excerpts below illustrate this function.

Mr. B1-13: because (.) sometime (.) we sleepover at her place .

Mr. B1-13: and her family know me .

Mr. B1-13: (..) so [DM] we &-uhh (.) drink and laugh .

Mr. B1-13: and nothing (.) everything .

(Mr. B1-13)

Excerpt 11

Ms. B1-29: I think she +/ .

Ms. B1-29: +/ oh (.) she &emm (.) we [/] we both in the same team .

Ms. B1-29: and (.) we play together .

INT: you mean PE lesson ?

Ms. B1-29: oh yes .

Ms. B1-29: we play volleyball (.) together .

INT: oh ok .

Ms. B1-29: yes .

Ms. B1-29: but (..) …+/ .

Ms. B1-29: so [DM] we become friend .

(Ms. B1-29)

Excerpt 12

In excerpt 11, the preceding context before *so* indicates that the participant often hangs out with his best friend at her house. It is overtly to see that *so* marks a result that they end up drinking, laughing and probably doing everything else together. However, it should be noted that the result illustrated above cannot be inferred from the preceding context directly, in which "sleepover" and "her family know me" cannot deduce the actions of "drink", "laugh" and "everything". It is therefore reasonable to conclude that *so* in this context only serves as a discourse marker to mark a result.

In the same condition, it can be interpreted that "we become friend" preceded by *so* marks a result with which the preceding context before *so* is not necessarily associated in excerpt 12. In other words, the fact that two people play volleyball in a team does not necessarily lead to the result of being friends together. Hence, *so* in this context serves as a discourse marker to mark a result.

A transition sequentially

In conversation, as illustrated by the previous research (Müller, 2005; Buysse, 2012), *so* yielded this function in a context where a short story that comprises a series of utterances consecutively was spelled out by the speaker, resulting the use of *so* in sequential. It is therefore conceivable that the use of the discourse marker *so*, under this circumstance, appears to be also consecutive. The excerpt below illustrates this function.

> Mr. B1-02: so [DM] (..) I want to study here because in Thailand is the best one .
>
> INT: ok .
>
> Mr. B1-02: yes so [DM] it's my dream .
>
> INT: ok
>
> Mr. B1-02: so [DM] when I achieve my goal I study here .

Mr. B1-02: so [DM] (..) <I want> [//] I like to attend in every activity .

(Mr. B1-02)

Excerpt 13

In this short context extracted from the original conversation, four consecutive utterances are preceded by four discourse marker *so*, consecutively. This phenomenon is in line with the demonstration by previous research (Müller, 2005), leading to the fact that the participant simply adopts *so* to elicit four consecutive utterances for the purpose of transition sequentially in that the utterances can be produced continuously.

Marking a start of a new narration

It is relatively simple to identify this function because *so* appears at the very beginning of a new utterance to lead a new topic or a new direction in conversation (Buysse, 2012). Both excerpts below illustrate this function.

Mr. B1-02: so [DM] if I go abroad and I come back .

Mr. B1-02: and if I have a chance .

Mr. B1-02: I want to change .

Mr. B1-02: I want to make change to Thailand .

(Mr. B1-02)

Excerpt 14

Ms. B1-12: so [DM] I have this best friend .

Ms. B1-12: she [/] she the same type with me .

(Ms. B1-12)

Excerpt 15

The use of the discourse marker *so* in both excerpts above suggests that the participants start a new topic or a new direction in conversation. It is also noticed that all the utterances following *so* are concerned with the same topic, i.e., to change Thailand in excerpt 14; an introduction of a best friend in excerpt 15. It is hence construed that *so*, in this condition, marks a new start of a narration.

Topic shift back

A subtle observation by the previous research (Tagliamonte, 2005; Müller, 2005; Buysse, 2012; Blakemore, 2014) discovered that this function should be identified in a context where the topic of the utterance preceded by *so* was illustrated in an earlier context that was followed by several other utterances referring to either the details of the same topic or some other topics. The following excerpt illustrates this function.

Mr. B1-10: (…) I'm (.) quite introvert .

INT: ok .

Mr. B1-10: yeah (.) I think everyone can see .

INT: <laugh> .

Mr. B1-10: I usually (..) stuck at home .

Mr. B1-10: (.) stuck at school .

Mr. B1-10: and (.) read .

Mr. B1-10: &-emm (..) yeah (.) do my thing .

Mr. B1-10: like play game .

Mr. B1-10: so [DM] I'm quite introvert .

(Mr. B1-10)

Excerpt 16

As stated above, it is clear to observe that the statement "I'm quite introvert" preceded by *so* has been addressed earlier at the beginning of this context before

another series of utterances is inserted that partly concerns some detailed description of being introvert narrated by the participant. It is thus simple to recognize that the discourse marker *so* in this context serves a topic-shift-back function.

A point worthy of note concerning this function is that this function can be identified only in a premier condition where a similar statement to the topic of the utterance preceded by *so* is addressed in an earlier context. Otherwise, the use of *so* should be interpreted by other functions.

Leading a summary

There are only two cases of using *so* as this function in BTELSEC, as illustrated by the following excerpt.

> Ms. B1-09: I don't (.) know .
>
> Ms. B1-09: she's not tall .
>
> Ms. B1-09: not &-uhh (.) just medium .
>
> Ms. B1-09: and she very kind .
>
> Ms. B1-09: and she [/] she [/] she like chat .
>
> INT: ok .
>
> Ms. B1-09: so [DM] (..) a good person (.) with good look (.) I think [DM].
>
> (Ms. B1-09)

Excerpt 17

This function can be identified when the utterance preceded by *so* is a summary of a series of utterances in the preceding context (Müller, 2005; Tagliamonte, 2005). The participant in excerpt 17 summarizes that her friend is a good-looking person with a good personality based on the details given by the utterances altogether in the preceding context, including the information of "not tall", "medium", "very kind" and

"like chat". It should be noted that the utterance preceded by *so* must be a summary of several utterances in an earlier context rather than just one utterance.

Marking opinion/feeling/stance

It was attested that this function existed commonly by different discourse markers, such as *so*, *I think*, *you know*, etc. (Brinton, 2008). It was also illustrated that the interpersonal functions were subjective functions (Brinton, 2008), including expressing responses, reactions, attitudes, etc. (Bazzanalla, 2006; Brinton, 2008; Jucker, 2002). Hence, marking the speaker's opinion, feeling or stance becomes a significant interpersonal function shared by many discourse markers.

As such, it is not surprising to see the high frequent use of this function by both native English speakers and non-native English speakers. Besides, it is also relatively easy to identify in that the utterance preceded by *so* indicates the opinion, the feeling or the stance of the speaker, as illustrated by the excerpts below.

> Ms. B1-04: <my mom> [//] (.) she kind and she good (.) to help me .
>
> Ms. B1-04: with my [/] my study .
>
> Ms. B1-04: so [DM] (.) I'm happy .
>
> Ms. B1-04: yeah .
>
> (Ms. B1-04)

Excerpt 18

> Mr. B1-23: I feel (..) many people don't have job .
>
> Mr. B1-23: I don't know the future .
>
> Mr. B1-23: so [DM] <I don't> [//] not sure about the future now .
>
> (Mr. B1-23)

Excerpt 19

The participants in both excerpts above express their feelings towards something they discussed in the preceding context, respectively. The first participant feels happy because her mom helps her a lot during her years of growing up, while the second participant expresses his worries about his future because of the special situation nowadays. It is noticed that both statements of the speakers' feelings are preceded by the discourse marker *so*. Given that the majority of the instances involving *so* as an interpersonal function is discovered to be this one as such, it is hence conceivable that Thai B1-level EFL learners perform the discourse marker *so* limited in one type of interpersonal function as marking their opinions, feelings or stances.

Prompting the hearer to speak/an implied result

In conversation, the speaker may encounter a situation in which a series of utterances have been finished narrating while it is the hearer's turn to lead the conversation. It is thus assumed that the speaker will trigger some signal to the hearer for the turn-taking. Under the circumstances, the discourse marker *so* is placed right after a long series of utterances to prompt the hearer to speak (Bazzanalla, 2006; Buysse, 2012). It is interesting to notice that there is one case in BTELSEC that involves the discourse marker *so* as this function. This case is illustrated below.

Mr. B1-02: when I in high school .

Mr. B1-02: I always have argument .

Mr. B1-02: but when I grow up I know <what they want to> [/] (.) .

Mr. B1-02: what they want to tell me .

Mr. B1-02: but actually I thought .

Mr. B1-02: ok I know that ok .

Mr. B1-02: so [DM] (.) +/

INT: <laugh> sure ok .

(Mr. B1-02)

Excerpt 20

In this context, the participant tells the hearer that he used to argue with his parents when he was a child and a teenager, whereas he can understand them as an adult now. He spells out a lexicon *so* with a stretch of the vowel [oʊ] in it without more words after a long monologue. It is thereby associated with the fact that the speaker has finished the leading role in conversation and intends to give it back to the hearer. The result is that the hearer reaches out the signal and starts to talk again.

A point worth mentioning in this case is that the speaker stretches the vowel [oʊ] in *so* as a phonetic sign to trigger the signal.

Hold the floor

In contrast to the function stated above, it is also observed another situation in which the speaker intends to remain the leading role in conversation even after having illustrated a long monologue. There is also only one case of using *so* as this function in BTELSEC, as illustrated below.

> INT: why you wanna stay in Bangkok ?
>
> Mr. B1-23: &-uhh (..) it's a big city .
>
> Mr. B1-23: many opportunities .
>
> Mr. B1-23: many company here .
>
> Mr. B1-23: and &-emm (..) I (.) think [DM] (.) many Thai people want to be here .
>
> INT: ok .
>
> INT: I guess so .
>
> Mr. B1-23: but (..) now (.) covid 19 .
>
> Mr. B1-23: so [DM] (..) .
>
> Mr. B1-23: I'm not sure .

Mr. B1-23: I just guess .

Mr. B1-23: I think [DM] (.) not many people now .

(Mr. B1-23)

Excerpt 21

In this context, a series of utterances have been spelled out before *so*, discussing the reason why many Thai people choose to live in a bigger city. It is noticed that a new information, mentioning Covid-19, is inserted before *so*. After producing *so* with a long pause, the speaker starts to discuss more information continuously. Considering that the speaker continues the conversation with another series of utterances after a long monologue in the preceding context, the discourse marker *so* lends support to hold the floor for the speaker so that he remains the leading role in conversation.

To summarize, this research found that Thai B1-level EFL learners used *so* as a discourse marker in both textual functions and interpersonal functions. Among the textual functions, marking a result or consequence was used the most. In the interpersonal functions, Thai B1-level EFL learners used *so* to express their opinion, feeling and stance the most. Compared to the textual functions, B1-level learners used *so* in a much lower frequency.

4.4.1.3 Pragmatic functions of *you know* as a discourse marker

Based on the data of BTELSEC, it was found that Thai B1-level EFL learners performed one textual function of the discourse marker *you know* with one case only in addition to four types of interpersonal functions of it, as is summarized in Table 30 and Table 31 below.

Table 30:Textual functions of you know as a discourse marker used by Thai B1-level

EFL learners

Textual functions of *you know* as a discourse marker	Number of cases in the raw data
Searching for the right word	1

Table 31: Interpersonal functions of you know as a discourse marker used by Thai

B1-level EFL learners

Interpersonal functions of *you know* as a discourse marker	Number of cases in the raw data
To acquire an understanding/an involvement from the hearer	16
An implicature of shared or common knowledge	3
An implicature of personal opinion/feeling/stance	1
Emphasis/repetition	1

As is shown in both tables above, it can be seen that the interpersonal function, namely, to acquire an understanding or an involvement from the hearer, became the predominant function performed by Thai B1-level EFL learners, whereas all the other functions were seldom performed.

The details of the identification of each function are presented as follows.

Searching for the right word

The only case that belongs to the textual function of *you know* is shown below.

Mr. B1-16: and (.) because they know (.) +…

Mr. B1-16: just (…) .

Mr. B1-16: you know [DM] (..) just speech .

Mr. B1-16: not action .

(Mr. B1-16)

Excerpt 22

It is observed at the moment when the participant evidently stammers out the statement illustrated as "just speech not action" preceded by a discourse marker *you know* in this context where the stammer starts from the beginning, with usual longer pauses after some unfinished utterances, i.e., "because they know", "just", "you know". It is thus considered that the participant is struggling to look for a right word in the next utterance. This finding is also aligned with previous research (Müller, 2005; Vanda & Peter, 2011) in which non-native English speakers perform this function to search the right words in conversation.

To acquire an understanding/an involvement from the hearer

This function was commonly discovered in conversation in reference to the sustainable maintenance of the interpersonal relationship between the speaker and other interlocutors (Brinton, 1996; Landgrebe, 2012). It was argued that the speaker, intentionally more or less, employed *you know* among utterances to acquire an understanding from the hearer at the moment, or to achieve an involvement as the conversation continues (Landgrebe; 2012). As illustrated by previous research (Brinton, 1996; Landgrebe, 2012; Müller, 2005; Vanda & Peter, 2011), to identify this function, the speaker shared information that related to the topic of the context at the moment when the discourse marker *you know* may be spelled out either before or after the information. Two excerpts below illustrate this function.

Mr. B1-16: <they> [/] they teach us simple as simple mathematics .

vas simple English .

Mr. B1-16: but (.) entrance exam require more than that .

Mr. B1-16: you know [DM] .

(Mr. B1-16)

Excerpt 23

Mr. B1-16: but in my opinion (.) <it's too easy> [/] it's too easy .

Mr. B1-16: but it's not enough .

Mr. B1-16: if you want to study in this university .

Mr. B1-16: you know [DM] .

(Mr. B1-16)

Excerpt 24

As is seen in both excerpts, the participant shares information based on the topic that both the speaker and the hearer are discussing in the context at the moment, i.e., the university entrance examination of math requires more than what the students learn at school in excerpt 23; the students who want to apply for this university should study more in excerpt 24. It is noticed that both *you know* in the excerpts above are positioned at the end of the provided information, resulting in an interaction from the speaker's perspective on his own initiative. It is hence concluded that the speaker acquires an understanding of the information provided in the prior utterance from the hearer.

An implicature of shared or common knowledge

This function is performed when the speaker indicates a shared or common knowledge that is universal to all the interlocutors or all human beings (Müller, 2005; Vanda & Peter, 2011). To fulfill this function, it is observed that the utterance should be the universal information preceded by a discourse marker *you know*, as illustrated by the excerpt below.

169

Mr. B1-16: [<] you know [DM] <every ever-> [//] .

Mr. B1-16: everyone is come from different faculty [>] .

INT: [<] yes of course .

(Mr. B1-16)

Excerpt 25

It naturally occurs to the speaker and the hearer that the members of a student union in a university must come from different faculties. Hence, to explain more events regarding this common information in the following context, the participant inserts a discourse marker *you know* to indicate it to the hearer, while the hearer also responds to this shared information by stating the utterance "yes of course".

A point worth explaining further concerning this function to the prior one should be noted here. It may occur, to a certain level, that the information "entrance exam require more than that" in excerpt 23, analyzed in the previous function, can be treated as a piece of common knowledge to both the interlocutors or even a large amount of people, so to speak. However, it is the different positions of the discourse marker *you know* spelled out by the speaker that determine the different types of functions, as demonstrated by the findings of the previous research (Müller, 2005; Vanda & Peter, 2011). The position of the discourse marker *you know* at the end of an utterance in this case accounts for acquiring an understanding or an involvement from the hearer, whereas the position that is at the start of the utterance leads to a common knowledge under this function category.

An implicature of personal opinion/feeling/stance

The same as other discourse markers illustrated earlier, the discourse marker *you know* is also discovered as this function to indicate personal opinion, feeling or stance (House, 2009; Landgrebe, 2012), as is illustrated by the only case found below.

Mr. B1-16: you know [DM] (.) for me it's bad .

Mr. B1-16: it sounds weird .

INT: yeah it's weird .

(Mr. B1-16)

Excerpt 26

The case above shows that the participant takes a stance to some information in the preceding context with the statement "for me it's bad" preceded by the discourse marker *you know*.

Emphasis/repetition

It was found that native English speakers may employ *you know* with respect to the emphasis or repetition of some information (Müller, 2005), in which the information was delivered in reiteration with a certain strengthened tone, as illustrated by the only case found in the data.

INT: so [DM] (.) you still have any argument right now ?

Mr. B1-16: (.) &-emm (.) some ideas .

Mr. B1-16: but you know [DM] (.) sometimes just sometimes .

INT: ok .

(Mr. B1-16)

Excerpt 27

To emphasize the information shown in the excerpt above, it is evident to see that the participant spells out the word "sometimes" twice by using a "just" to strengthen the tone for the second time. It embodies the fact that the word "sometimes" is the information emphasized by the speaker.

In summary, it was found that seldom cases were identified to the use of the discourse marker *you know*. It should be noted that all the cases come from the same participant. It is hence conceivable that Thai B1-level EFL learners were lack the awareness or the knowledge to use *you know* as a discourse marker.

4.4.1.4 Pragmatic functions of *I think* as a discourse marker

As the result showed previously, it is the only discourse marker that was overused by Thai B1-level EFL learners under investigation in this research. Two pragmatic functions were identified, which comprised one textual function and one interpersonal function, as summarized in Table 32 below.

Table 32: Pragmatic functions of I think as a discourse marker used by Thai B1-level EFL learners

Pragmatic functions of *I think* as a discourse marker	Number of cases in the raw data
An implicature of personal opinion/feeling/stance	204
Searching for the right word	35

As the table above shows, it is observed that the interpersonal function, namely, an implicature of personal opinion, feeling or stance, dominated the use of the discourse marker *I think* by Thai B1-level EFL learners, while searching for the right word, as a textual function of *I think*, was performed much less frequently.

The details of the identification of both functions of the discourse marker *I think* are presented below.

An implicature of personal opinion/feeling/stance

It was argued by Aijmer (1997) in the early studies of the spoken discourse markers that the collocation *I think*, whose semantic meaning was to express the

speaker's thought as a NP subject to coordinate *that*-clauses, was frequently used by native English speakers in conversation, yet the semantic meaning of it was diminished in dialogic action. It was investigated by the previous research (Brinton, 2008, 2010; Dehe & Wichmann, 2010; Diessel & Tomasello, 2001) that the collocational pattern *I think*, detaching its role of being a subject-verb form to lead *that*-clauses, in addition to its unstressed feature, should not be treated as a NP subject at the beginning of an utterance, but a discourse marker to lead a comment clause as illustrated by Brinton (2008). In other words, without the necessity of the attachment of *that*-clauses with respect to *that*-deletion, the collocational pattern *I think* may be understood as a discourse marker, constituting the subclass of epistemic or evidential parentheticals (Brinton, 2008) to indicate the speaker's opinion, feeling or stance uttered in a comment clause. This is because, in this condition, this collocational pattern does not form a hierarchical syntactic relationship with its anchor clause (Brinton, 2008).

Based on the theoretical background, 204 instances are discovered by using the discourse marker *I think* as this function, as illustrated by the excerpts below.

Ms. B1-04: I think [DM] (.) there can have more transportation .

(Ms. B1-04)

Excerpt 28

Ms. B1-12: but (.) I'm not sure .

Ms. B1-12: I see sometime .

Ms. B1-12: (.) I think [DM] (.) not good .

(Ms. B1-12)

Excerpt 29

INT: how's your hometown ?

Mr. B1-24: &-emm (…) not the same .

Mr. B1-24: (.) I choose here (.) I think [DM] .

(Mr. B1-24)

Excerpt 30

As demonstrated by the three excerpts above, it is observed that the discourse marker *I think*, with *that*-deletion, longer pause surrounded it, and unstressed tone spelled out by the participants, detaches itself from the utterances followed by it, respectively. It can appear either at the beginning of an utterance, as shown in excerpt 28 and excerpt 29, or at the end of an utterance like in excerpt 30. Its function is to mark a personal's opinion, feeling or stance.

Searching for the right word

As noted earlier in the analysis of the same function of other discourse markers under investigation in this research, the frequent co-occurrence, as the combination of the discourse marker *I think* and some fillers, such as *uhh* or *emm*, evidently illustrates this function, as shown in the excerpts below.

Mr. B1-16: childhood (…) .

Mr. B1-16: I think [DM] I &-uhh (…) .

Mr. B1-16: I grow up with a happy family and good friend .

(Mr. B1-16)

Excerpt 31

INT: do you like beach ?

Ms. B1-27: yes I like beach .

Ms. B1-27: I like quiet (.) and like sea .

Ms. B1-27: I think [DM] &-uhh (..) beach is relax .

(Ms. B1-27)

Excerpt 32

It is simple to notice that in both contexts shown above, the participants employ the discourse marker *I think*, followed by the filler *uhh* and longer pauses, to search for the word they are about to utter.

One point worth mentioning here is that because of the impact of the original semantic meaning of the spoken discourse marker *I think*, it might be considered, in the excerpts above as well as other instances from the learner corpus, that the function, an implicature of personal opinion, feeling or stance, was used by the speakers at the same time. This phenomenon is, as has been illustrated at the beginning of this part, that more than one pragmatic function may exist in a specific function. Once again, this research tried to explore the main function of each spoken discourse marker in each context. Because of the linguistic features in the excerpts above, i.e., a longer or an unusual pause when inserting *I think*, along with different fillers, the function, searching for the right words, was regarded as the main function in the context.

In summary, Thai B1-level EFL learners overused *I think* in English conversation where mostly, *I think* was performed as an implicature to show their personal opinions, feelings or stance.

4.4.1.5 Pragmatic functions of *I mean* as a discourse marker

Based on the data of BTELSEC, Thai B1-level EFL learners performed the discourse marker *I mean* as its two textual functions and one interpersonal function, as summarized in Table 33 and Table 34 below.

175

Table 33: Textual functions of I mean as a discourse marker used by Thai B1-level EFL learners

Textual functions of *I mean* as a discourse marker	Number of cases in the raw data
More explanation/exemplification	4
Marking a repair	3

Table 34: Interpersonal functions of I mean as a discourse marker used by Thai B1-level EFL learners

Interpersonal functions of *I mean* as a discourse marker	Number of cases in the raw data
An implicature of personal opinion/feeling/stance	1

As the two tables shown above, it can be seen that Thai B1-level EFL learners rarely performed the discourse marker *I mean*. The research presents the details as follows.

More explanation/exemplification

Like the other two multi-word discourse markers under investigation in this research, the NP form, with respect to the subject *I* and the verb *mean*, derives its semantic meaning of leading *that*-clauses into a collocational pattern that denotes the relationship of the interlocutors in conversation, resulting in the attention to its pragmatic functions in conversational interaction (Brinton, 2010; Vanda & Peter, 2011). It was argued that the zero complimentizer of *I mean*, due to *that*-deletion in dialogic action, was seen as the derivation of its role as a discourse marker, in addition to a reference to a significant principle regarding the identification of discourse markers (Brinton, 2008, 2010).

As also noted earlier, some functions of a discourse marker may be derived from its original semantic meaning. It was found that this function was associated with the semantic meaning of the NP form of *I mean* that the use of it was to lead to an explanation from the speaker's perspective (Brinton, 2010). It is thereby perceived that the utterance preceded by the discourse marker *I mean* must provide an explanation or exemplification to a prior utterance in the preceding context, as illustrated by the excerpts below.

INT: how did you feel about your university life ?

Mr. B1-11: &-emm (…) I think [DM] (.) it is really friend .

Mr. B1-11: I mean [DM] (.) it is not different in (.) making friends .

INT: ok .

(Mr. B1-11)

Excerpt 33

Mr. B1-11: and (…) everything is so expensive .

Mr. B1-11: &-uhh I mean [DM] compare to other country .

Mr. B1-11: I mean [DM] in my country (..) rice is thirty Baht (..) .

INT: uh-huhh .

Mr. B1-11: but in Bangkok is forty or fifty Baht .

(Mr. B1-11)

Excerpt 34

As demonstrated in excerpt 33, the utterance "it is not different in making friends" preceded by the pattern *I mean* provides a further explanation of the prior utterance "it is really friend", which may be regarded as an utterance including a grammatical error, resulting in a difficult understanding from the hearer's point of view. It is therefore

177

delighted to see that the participant continues an explanation closely to it by performing a discourse marker *I mean* between the two utterances.

It is interesting to notice that the discourse marker *I mean* in excerpt 34 is performed twice consecutively under this same function, wherein the utterance "compare to other country" preceded by the first discourse marker *I mean* provides an additional explanation to the prior utterance "everything is so expensive", while the utterance starting "in my country" preceded by the second discourse marker *I mean* supplies an exemplification in detail.

Marking a repair

As the name of this function suggests, the speaker corrects a prior utterance partly after inserting the discourse marker I mean. It is assumed that the speaker realizes a mistake he or she utters in the prior utterance, as illustrated by the excerpt below.

> Mr. B1-11: (…) &-uhh (…) I tend to ask for extra money during the month .
> Mr. B1-11: I mean (…) I tend to ask for extra money from my dad .
> INT: ohh ok .
> (Mr. B1-11)
> **Excerpt 35**

The key to identifying this function is to examine whether there is a repair in the utterance preceded by *I mean* compared to the prior utterance. As demonstrated in the excerpt above, the information "during the month" in the prior utterance is substituted by a new one "from my dad". It is assumed that the participant wants to deliver the information "from my dad" rather than the information "during the month". It should be noted concerning this function that the identification of this function should be

based on the fact that only one part of the prior utterance, such as a word or a short expression, is repaired. Otherwise, other functions must be considered.

An implicature of personal opinion/feeling/stance

Only one interpersonal function with one case is found in BTELSEC, as illustrated as follows

Mr. B1-11: and he likes to play some games with me .

INT: ok .

Mr. B1-11: <and and> and I help to learn and study until six years .

Mr. B1-11: from high school .

Mr. B1-11: I mean [DM] &-uhh (…) nowadays that is something .

(Mr. B1-11)

Excerpt 36

It is evidently recognizable in the excerpt above that the participant illustrates his stance on his six-year friendship by stating the utterance "nowadays that is something".

To summarize, it was found that Thai B1-level EFL learners rarely performed the discourse marker *I mean*, as all the cases found in BTELSEC were produced from one participant.

4.4.2 Pragmatic functions of each spoken discourse markers used by Thai C1-level EFL learners

4.4.2.1 Pragmatic functions of *like* as a discourse marker

It was found that Thai C1-level EFL learners used the same six pragmatic functions as Thai B1-level EFL learners did. Table 35 shows the functions of the discourse marker *like* used by Thai C1-level EFL learners.

Table 35: Textual function of like as a discourse marker used by Thai C1-level EFL learners

Textual function of *like* as a discourse marker	Number of cases
Marking a focus on new information	179
To exemplify	104
As an approximator	65
To have a quotation	19
More information/explanation	13
Searching for the following words	10

As shown in Table 35, it was observed that Thai C1-level EFL learners used *like* as a discourse marker in a relatively different way with B1-level learners. Unlike B1-level learners, C1-level learners used the discourse marker *like* as the function of marking a focus on new information the most, followed by the function named as to exemplify. A huge difference is that C1-level learners used the discourse marker *like* to be an approximator with 65 cases, whereas only one case was found in B1-level learners. On the contrary, to search for the following words was relatively used in a low frequency by C1-level learners.

The details of each function are shown as follows.

Marking a focus on new information

Based on the data, it seems that C1-level learners had a preference to use this function. Excerpt 37 and excerpt 38 illustrate this point.

Mr. C1-01: (.) and somehow (.) I (.) felt (.) I'm uncomfortable with them .

Mr. C1-01: because (.) like [DM] (.) they are super rich .

(Mr. C1-01)

Excerpt 37

Mr. C1-06: I didn't feel any pressure at all .

Mr. C1-06: I can be me .

Mr. C1-06: I can be like [DM] super me .

(Mr. C1-06)

Excerpt 38

It is worth mentioning that out of 179 instances, 113 of them explicitly included a strengthened tone on the emphasized part spelled out by the speakers after the discourse marker *like* is given, such as in the two excerpts above.

The participant in excerpt 37 provides the reason why he feels uncomfortable with some of his classmates. While he delivers the reason, he inserts the discourse marker *like* between the word "because" and the reason part "they are super rich". The speaker deliberately pauses a while before and after the discourse marker *like* and strengthens his tone on the reason part. The participant in excerpt 38 inserts the discourse marker *like* before he strengthens his tone on the word "super" so that he emphasizes the fact that he can totally be himself in front of his friends.

Based on the detailed analysis above, it reflects the phenomenon that Thai C1-level EFL learners performed this function more proficiently than Thai B1-level EFL learners in that C1-level learners used this function more frequently wherein many of them strengthened their tones on the new information at the same time.

To exemplify

Thai C1-level EFL learners performed this function the same as Thai B1-level EFL learners did, as illustrated in the excerpt below.

Ms. C1-12: she's (..) not very careful (.) .

Ms. C1-12: but sometimes I think +/ .

Ms. C1-12: +/ (.) but not serious .

181

Ms. C1-12: like [DM] sometime she don't tell me .

Ms. C1-12: and she's [/] <she she> (.) she take my things .

Ms. C1-12: but she didn't tell me .

(Ms. C1-12)

Excerpt 39

In excerpt 39, it can be interpreted from the preceding context before the discourse marker *like* that the personality of the speaker's friend can be seen as insignificant carelessness. To achieve a better mutual understanding with the hearer, the speaker spells out a series of utterances led by *like*, manifesting the concrete behavior of significant carelessness, i.e., don't tell; take my things; didn't tell me. It is thereby more explicit to identify the meaning of the preceding context before *like* with the exemplifications of her friend's actions.

As an approximator

As has been stated above, Thai C1-level EFL learners used the discourse marker *like* as an approximator much more often than B1-level learners. It was stated by the previous research (Diskin, 2017; Müller, 2005) that this function was performed when the speaker signaled an approximate number, as illustrated in the two excerpts below.

Mr. C1-01: for example (..) <the the> the English section course .

Mr. C1-01: they have like [DM] ten sections .

Mr. C1-01: I don't remember .

(Mr. C1-01)

Excerpt 40

Mr. C1-22: because when I look back .

Mr. C1-22: I just met my friend like [DM] (.) two weeks ago .

182

(Mr. C1-22)

Excerpt 41

It is clear to address the pattern "*like* + a number" from both excerpts, wherein this function is rendered by it (Tagliamonte, 2005). The participant in excerpt 40 is not sure if the English course he refers to has exact ten sections. He thereby inserts the discourse marker *like* before the number "ten" to signal that the number "ten" may not be accurate. The participant in excerpt 41 also performs the discourse marker *like* in front of the number "two" to signal that he met his friend approximately two weeks ago.

To have a quotation

19 cases were found where Thai C1-level EFL learners used this function, as shown in the excerpt below.

> Ms. C1-28: that moment makes me wanna scream .
>
> INT: <laugh> .
>
> Ms. C1-28: like [DM] (.) .
>
> Ms. C1-28: oh (.) I have a question .
>
> Ms. C1-28: teacher (.) can you say it again ?
>
> (Ms. C1-28).

Excerpt 42

It is interesting to note here that 8 out of 19 cases in which Thai C1-level EFL learners performed this function was adopted an imitation of the sound in a way that the quotation was produced in the original scene, such as in the excerpt above. With the actual vivid sound imitation, the participant in excerpt 42 addresses the quotations

more vivaciously. This phenomenon does not happen when Thai B1-level EFL learners used this function.

More information/explanation

This function was relatively used in a low frequency by Thai C1-level EFL learners compared to other functions they used. The excerpt below illustrates this function.

> Mr. C1-09: but actually emm (..) those not knowledge .
>
> Mr. C1-09: we will [/] will have test (.) for that knowledge .
>
> Mr. C1-09: but knowledge (.) like [DM] (.) general knowledge .
>
> Mr. C1-09: like [DM] many things in life .
>
> (Mr. C1-09)

Excerpt 43

It is noted that the participant in the excerpt above uses this function twice consecutively in a short context in which each provided information or explanation supplements the prior information step by step sustainably. The first *like* in excerpt 43 preceding the more information "general" is performed to illustrate the type of "knowledge" followed by it, while the second provided explanation "many things in life" preceded by the second *like* is addressed to explain the concept of "general knowledge". As stated in the analysis of this function used by Thai B1-level EFL learners, each provided information or explanation in this excerpt is short and simple, with one more word or a short expression.

Searching for the following words

Thai C1-level EFL learners used this function the same as Thai B1-level EFL learners did. It is observed that the discourse marker *like* co-occurred with the fillers in each case, i.e. *uhh*, *uhmm*, as illustrated in the excerpt below.

Ms. C1-12: she's trying something .

Ms. C1-12: but (.) she told me (.) now is hard to (.) get a ticket .

Ms. C1-12: and (.) something like [DM] (..) &-uhmm (..) .

Ms. C1-12: they sort of closed the airport .

(Ms. C1-12)

Excerpt 44

In excerpt 44, the participant tries to express the difficulty of buying a flight ticket to come back to Thailand. After the discourse marker *like*, there is a relatively long pause with the filler *uhmm*. It can be interpreted that she uses *like* as this function to think of what she is going to say next.

To summarize, Thai C1-level EFL learners used the discourse marker *like* in total six textual functions, in which marking the focus on the new information and to exemplify were the two predominant functions. Furthermore, the function named as an approximator was used much more often than Thai B1-level EFL learners.

4.4.2.2 Pragmatic functions of *so* as a discourse marker

Similar to the situation of how Thai B1-level EFL learners used it, Thai C1-level EFL learners also used *so* as a discourse marker in both textual functions and interpersonal functions. Based on the original data, it was found that Thai C1-level EFL learners used six textual functions and four interpersonal functions. Table 36 and Table 37 summarize the functions that Thai C1-level EFL learners used in English conversation.

Table 36: Textual functions of so as a discourse marker used by Thai C1-level EFL learners

Textual functions of *so* as a discourse marker	Number of cases in the raw data

Marking a result/consequence	57
Marking a start of a narration	25
Leading a summary	17
A transition sequentially	3
Topic shift back	2
More explanation	1

Table 37: Interpersonal functions of so as a discourse marker used by Thai C1-level

EFL learners

Interpersonal functions of *so* as a discourse marker	Number of cases in the raw data
Marking opinion/feeling/stance	32
Prompting the hearer to speak/an implied result	5
hold the floor	5
Marking a question	3

The same as the case of Thai B1-level EFL learners, for the textual functions of the discourse marker *so*, marking a result or a consequence owned the predominance. On the contrary, two other functions, namely, marking a start of a narration and leading a summary, which were not the leading textual functions used by B1-level learners, became the main textual functions used by C1-level learners. Besides, a new textual function, namely providing more explanation, which was not performed by B1-level learners, was identified in CTELSEC with one case only.

On the other hand, the frequency distribution of using the interpersonal functions by Thai C1-level EFL learners was almost the same as Thai B1-level EFL learners. The function that was mostly used by C1-level learners was marking opinions,

feelings or stances. One difference is that a new interpersonal function was discovered in CTELSEC, which was marking a question or a request.

The details of each function that was performed in CTELSEC are presented as follows.

Marking a result/consequence

It was found that Thai C1-level EFL learners also used the discourse marker *so* to mark a result or a consequence at a high frequency as B1-level learners did. Two excerpts extracted from CTELSEC are presented below.

Mr. C1-16: yes (.) and I just live by the beach when I grow up .

Mr. C1-16: so [DM] (.) I went to see mountains .

(Mr. C1-16)

Excerpt 45

Ms. C1-08: but everyone has their first time .

Ms. C1-08: so [DM] (.) I (.) just do what I do .

(Ms. C1-08)

Excerpt 46

As stated in the previous part, the utterance preceded by *so* in the context only serves as a result or a consequence of the prior utterance, wherein the two utterances do not have the direct cause-and-effect logic. The participant in excerpt 45 states the utterance that he goes to the mountains as a result of the fact that he has been living by the sea throughout his childhood. However, it is reasonable to notice that the prior utterance before *so* is an unnecessary but sufficient condition for the result. In other words, the participant who grew up by the sea cannot directly lead to the result that he goes to the mountains.

187

As demonstrated by excerpt 46, the participant shares a common sense that everyone has the first time to do something that is associated with the fact that it is impossible for everyone to fully prepare or know exactly what to do the first time. It is thereby illustrated by the participant, after the discourse marker *so*, that as a result, she simply does everything the way she is supposed to do. By inserting the discourse marker *so*, the speaker leads a result of the prior utterance.

Marking a start of a narration

It should be noted that to make this function take place in conversation, it is essential for the speaker to lead a new start of a narration on his or her own initiative. In other words, the participant involved in this research should be willing to narrate more on some topic. It is therefore assumed that Thai C1-level EFL learners performed this function in a high frequency is because they were able to perform better in spoken English. The excerpt below illustrates this function.

> Ms. C1-04: so [DM] (.) when we went to interview the prisoner .
>
> Ms. C1-04: he's strict about everything .
>
> Ms. C1-04: we need to take note .
>
> (Ms. C1-04)

Excerpt 47

The participant in the excerpt above leads a narration by performing the discourse marker *so*, describing the experience of her going to interview a prisoner with her advisor in detail. With a better language performance pointed out above, it is thus conceivable that Thai C1-level EFL learners use this function more frequently than Thai B1-level EFL learners.

Leading a summary

It was observed that Thai C1-level EFL learners performed this function more frequently than Thai B1-level EFL learners. One of the excerpts from the data is shown below.

Mr. C1-01: I mean [DM] (.) we were so good at studying .

Mr. C1-01: and (.) we know (.) how to make good grades .

Mr. C1-01: good paper (.) good (.) everything .

Mr. C1-01: what the teacher wanted .

Mr. C1-01: so [DM] (.) we did it .

(Mr. C1-01)

Excerpt 48

The utterance after the discourse marker *so* is a whole summary of all the previous utterances, referring to the fact that they do everything well as required by the teachers, including studying hard, achieving good scores and writing good papers.

A point worth mentioning here is that it is also assumed that a higher proficiency of English-speaking ability of Thai C1-level EFL learners accounts for the high frequency of this function in that a series of utterances must be spelled out before a summary.

A transition sequentially

As illustrated by the data, it can be seen that Thai C1-level EFL learners seldom performed this function. One excerpt is shown below.

Ms. C1-20: [<] exactly like you said .

Ms. C1-20: so [DM] (.) the place I lived is in the [/] like [DM] a suburb .

Ms. C1-20: so [DM] (.) if you take a car .

Ms. C1-20: it will take you like [DM] 20 to 30 (.) minutes .

(Ms. C1-20)

The first *so* in the excerpt above is to mark a start of a narration. The participant starts a new narration on where she lived when she was abroad. In order to continue the narration, the participant adds another *so* as a discourse marker to connect another utterance that is associated with the same topic with more details. It is therefore considered that the second *so* in this context serves as a transition sequentially.

Topic shift back

Only two cases were found in CTELSEC, as shown below.

Mr. C1-01: but (.) we changed it .

Mr. C1-01: (.) we were suppose to take &-uhmm (.) .

Mr. C1-01: what was that (..) .

Mr. C1-01: British literature .

Mr. C1-01: because our names were in the first section .

Mr. C1-01: but we somehow got to change .

Mr. C1-01: and (.) I lied to the teacher .

Mr. C1-01: so [DM] (.) we changed the British literature .

(Mr. C1-01)

Excerpt 50

Ms. C1-10: so [DM] I prefer to go to north of Thailand to see mountain .

Ms. C1-10: it has many mountains there .

Ms. C1-10: I prefer to go alone sometime .

Ms. C1-10: or with my sister .

Ms. C1-10: if she also free .

INT: ok what about the beach ?

190

Ms. C1-10: &-emm (..) I also like it .

Ms. C1-10: but (.) maybe (.) because I come from Chon Buri .

Ms. C1-10: and (.) my family is near the beach .

INT: oh ok .

Ms. C1-10: yes (.) and (.) I just live by the beach when I grew up .

Ms. C1-10: so [DM] (.) maybe [//] that's why I prefer mountain than the beach .

(Ms. C1-10)

Excerpt 51

It is overtly to recognize in both excerpts that the topic of the utterance preceded by *so* has been illustrated in an earlier utterance. In excerpt 50, the statement "we changed the British literature" preceded by *so* has been uttered at the beginning of this context as "we changed it", while the utterance "I prefer mountain than beach" preceded by *so* has also been stated at the beginning of the context as "I prefer to go to north of Thailand to see mountain". The series of utterances between the two statements in both excerpts partly explains the details, i.e., how they changed the subject in excerpt 50 and a discussion of the preference of the beach or the mountain. Under this circumstance, it is hence construed that the discourse marker *so* in both contexts yields the topic-shift-back function.

More explanation

It was discovered that the speaker used *so* to lead more explanation to a prior utterance, wherein the explanation provided details or more information to assist the full understanding in terms of the hearer's perspective (Buysse, 2012; Müller, 2005). Only one case was found in the data and was illustrated below.

INT: what's the different part ?

191

Ms. C1-18: &-uhh (.) it has many changes .

Ms. C1-18: (.) so [DM] (.) one subject got one (.) verb form .

Ms. C1-18: and (..) with male and female form and stuff .

(Ms. C1-18)

Excerpt 52

It is observed that a prior utterance, namely "it has many changes", is stated before the discourse marker *so*, denoting the meaning in this context that Spanish, the foreign language that is studied by the participant, contains many morphological changes. To fulfill the full understanding from the hearer's perspective, the participant continues the conversation with two utterances preceded by *so*, presenting two pieces of explanation to the prior utterance. In this context, the function of the discourse marker *so* is therefore rendered by the explanation in detail of the "many changes" stated in the prior utterance. Hence, the discourse marker *so* in this context marks more explanation as such.

Marking opinion/feeling/stance

As the data illustrated, this function, like the situation of Thai B1-level EFL learners, is the predominant interpersonal function of the discourse marker *so* performed by Thai C1-level EFL learners. Two excerpts below illustrated this function.

Ms. C1-08: we have some patients and (.) my classmates .

Ms. C1-08: and (.) we need to do something &-emm (.) .

Ms. C1-08: like [DM] something professor asked us to do .

Ms. C1-08: but it's the first time .

Ms. C1-08: so [DM] actually I feel excited .

(Ms. C1-08)

Excerpt 53

Ms. C1-24: +/ (.) and she just stay at home .

Ms. C1-24: but she live in a big house .

Ms. C1-24: and with &-uhh (.) other students .

Ms. C1-24: people .

Ms. C1-24: so [DM] I guess she's ok .

(Ms. C1-24)

Excerpt 54

The utterances preceded by *so* in both excerpts above express the feeling and the opinion of the participants, respectively. It is simple to recognize this point by examining both utterances in which "I feel" and "I guess" are stated as expressions to spread out the feeling and the opinion of the participants, respectively.

Prompting the hearer to speak/an implied result

As discussed previously, the speaker leaves a discourse marker *so* at the end of a relatively long series of utterances to signal the hearer for the turn-taking. Two excerpts extracted from CTELSEC illustrate this function.

Ms. C1-05: I want to (.) &-uhh (.) .

Ms. C1-05: our [/] my parents plan to go to Australia this year .

Ms. C1-05: but Covid 19 .

Ms. C1-05: so (..) .

INT: oh I see .

INT: sorry to hear that .

(Ms. C1-05)

Excerpt 55

Ms. C1-19: but my home is not in the city centre .

Ms. C1-19: and (.) I'm lazy .

Ms. C1-19: I don't go out often .

Ms. C1-19: so (..) .

INT: oh ok I get it .

(Ms. C1-19)

Excerpt 56

The participant in excerpt 55 produces a *so* after illustrating the fact that her family's plan to visit Australia has been interfered by Covid-19, while the participant in excerpt 56 inserts a *so* after stating the reasons why she did not have many activities. It is noticed that both participants cease their leading role in conversation with a long pause and no more utterance after the discourse marker *so*. It is also conceivable that the hearer continues the conversation in both excerpts, reflecting that the hearer receives the signal from the previous speaker and becomes the current speaker to continue the conversation. Hence, both *so* prompt the hearer to speak in this condition.

Hold the floor

As the data shows, Thai C1-level EFL learners also seldom used this function. An excerpt illustrates this function.

Ms. C1-04: actually I wanna go abroad to see .

Ms. C1-04: but my mom will have a problem with me .

Ms. C1-04: so (.) .

Ms. C1-04: we will see .

Ms. C1-04: hope she's ok in the future .

194

(Ms. C1-04)

Excerpt 57

Although the participant in the excerpt above inserts a discourse marker *so* after some illustrations, she continues the conversation with more utterances after *so* with a longer pause. It can be concluded that the discourse marker *so* in this context serves the function to hold the floor.

Marking a question

It is in line with the previous research (Brinton, 2008; Jucker, 2002; Müller, 2005) that some discourse markers mark a question or a request, including *so*. Unlike the situation of Thai B1-level EFL learners, it is discovered that Thai C1-level EFL learners performed the discourse marker *so* in this function. Three cases, found in CTELSEC, all reflect that the discourse marker *so* was used to mark a question, whereas marking a request was not discovered in the data. Two excerpts from the data illustrate this function.

Ms. C1-04: so [DM] (.) you want to teach Chinese ?

(Ms. C1-04)

Excerpt 58

Ms. C1-15: so [DM] (.) doing research is difficult ?

Ms. C1-15: &-emm (.) it depends on people .

(Ms. C1-15)

Excerpt 59

The two excerpts above show that Thai C1-level EFL learners used the discourse marker *so* to mark a question in conversation to signal to the hearer that they want to

know some information. It should be noted that like the functions, namely, marking a start of a narration and leading a summary, it is required the participants to ask a question or a request on their own initiative to make this function take place in conversation. It is thus assumed that Thai C1-level EFL learners have this function performed whereas no case was found from Thai B1-level EFL learners due to a higher speaking proficiency ability of the former and a lower speaking proficiency level of the latter.

To summarize, it was found that marking a result or a consequence was the predominant textual function of the discourse marker *so*, and marking an opinion, feeling or stance was the predominant interpersonal function of the discourse marker *so* performed by Thai C1-level EFL learners. Furthermore, they also performed two functions that were not discovered by Thai B1-level EFL learners: marking more explanation and marking a question.

4.4.2.3 Pragmatic functions of *well* as a discourse marker

As demonstrated by the data, it was found that Thai C1-level EFL learners performed *well* as a discourse marker for both textual functions and interpersonal functions, wherein two textual functions were performed and four interpersonal functions were performed, respectively. Table 38 and Table 39 summarize the functions that Thai C1-level EFL learners use in English conversation.

Table 38: Textual functions of well as a discourse marker used by Thai C1-level EFL learners

Textual functions of *well* as a discourse marker	Number of cases in the raw data
Marking a new topic	4
Searching for the right words	2

Table 39: Interpersonal functions of well as a discourse marker used by Thai C1-level EFL learners

Interpersonal functions of *well* as a discourse marker	Number of cases in the raw data
Marking a stance/feeling/opinion	7
Indicating an incomplete answer	4
Indicating an unexpected answer	4
Marking a dispreferred response	3

As illustrated by the data, although Thai C1-level EFL learners produced the discourse marker *well* in English conversation, the frequency was low and not many instances were found. Unlike the other two single-word discourse markers under investigation in this research, more instances were identified in the interpersonal functions of the discourse marker *well* than its textual functions, reflecting its main role in conversational interaction, in line with the finding of the previous research (Blakemore, 2014; Fung & Carter, 2007; Jucker, 2002; Lam, 2009; Müller, 2005; Norrick, 2001; Sakita, 2013), that the discourse marker *well* mainly accounted for the interpersonal interaction to structure the speaker's relationship to other interlocutors in conversation.

Among the six pragmatic functions discovered in CTELSEC, marking a stance, a feeling or an opinion led to its role as the predominant function, while other functions were involved in more or less two to four instances, respectively.

The details of the identification of each function are presented below.

Marking a new topic

It was found that in conversation, the speaker inserted the discourse marker *well* before a topic shifting, denoting either a condition of a topic shift back or a sign of a change to a new topic (Norrick, 2001). It is thereby comprehensible that there is a

discernible topic difference in both contexts surrounding the discourse marker *well*. The excerpt below illustrates this function.

> Ms. C1-13: always (.) the transportation is bad .
>
> Ms. C1-13: bothers me a lot .
>
> INT: <laugh> .
>
> Ms. C1-13: but now is getting (.) +/ .
>
> Ms. C1-13: +/ we have BTS now .
>
> INT: yeah of course .
>
> INT: but I notice the cars are still too many right ?
>
> Ms. C1-13: yeah I know .
>
> Ms. C1-13: the traffic jam always bothers me .
>
> INT: oh yeah ?
>
> Ms. C1-13: well [DM] (.) the pollution in Bangkok is bother too .
>
> Ms. C1-13: the air is not so clear now .
>
> (Ms. C1-13)
>
> **Excerpt 60**

Upon the examination of the excerpt above, it is firstly noticed that the topic of the context in this excerpt encompasses two problems in Bangkok: the poor transportation environment that consists of a long series of utterances before the discourse marker *well*; and the pollution problem in Bangkok appears to be after the discourse marker *well*. It is assumed that the participant perceives that it might be too sudden to attempt to have a new topic shifted immediately after the long discussion on the former one. In this condition, a *well* is inserted to elicit the new topic, namely the pollution problem in Bangkok, signaling the hearer to continue the conversation in this topic.

Searching for the right word

It was illustrated by the previous research (Jucker, 2002) that native English speakers used this function of the discourse marker *well* frequently in conversation. Like the similar function of the discourse marker *like* analyzed in the previous part, it was found that the discourse marker *well*, spelled out by the speaker for the use of this function, may co-occur with some fillers at the same time, i.e., *uhh*, *emm*, etc. (Fung & Carter, 2007; Jucker, 2002). This frequent co-occurrence, in the meantime, lends support to identifying this function in that it is a sign for the speaker to consider the words they are about to utter in mind. The excerpt below illustrates this function.

> INT: do you like online class ?
>
> Mr. C1-22: (.) not really .
>
> INT: <laugh> .
>
> INT: why ?
>
> Mr. C1-22: (.) well [DM] (..) &-uhh (.) first is (.) I <have to> [//] have a better internet .
>
> Mr. C1-22: I guess so .
>
> (Mr. C1-22)

Excerpt 61

It is recognizable in the excerpt above that the participant, at the moment of answering the prior question "why", stutters evidently, with the combination of several longer pauses in different lengths, repetition of some words, and the use of the filler *uhh* with which the discourse marker *well* associates, suggesting that he is thinking of how to answer the prior question. It is thus believed that the discourse marker *well* in this context serves the function of searching for the right word.

A point should be noted that although it was found that this function was used frequently by native English speakers stated above, while seldom cases were

199

identified in CTELSEC, which assumes that Thai C1-level EFL learners may lack the awareness of this function.

Marking a personal stance/feeling/opinion

Given that marking a personal stance, feeling or opinion is a main interpersonal function shared by different discourse markers in conversation (Aijmer, 2016; Brinton, 2008; Bolden, 2015), in addition to a major function of the discourse marker *well* supported by many studies (Blakemore, 2014; Fung & Carter, 2007; Lam, 2009; Müller, 2005; Norrick, 2001; Sakita, 2013), it is hence not surprising to see that Thai C1-level EFL learners performed this function in the highest frequency. Furthermore, the previous findings also revealed the idea that the generic use of the discourse marker *well* by native English speakers denoted various stances in different contextual environments, reflecting its intrinsic nature of stance-taking that, on the other hand, was inherent in dialogic activity emerging from the interaction between interlocutors (Sakita, 2013). It is thereby inferred, with no doubt, that this function should be widely used in English conversation. The two excerpts below illustrate this function.

Ms. C1-13: (.) and my dad like [DM] ask me to study economics .

Ms. C1-13: cause he thinks (.) it can easier to find a job .

Ms. C1-13: something like that .

INT: oh really ?

Ms. C1-13: well [DM] (.) it's his idea .

Ms. C1-13: anyway .

(Ms. C1-13)

Excerpt 62

Ms. C1-03: but it has dance in it .

Ms. C1-03: not just cheering .

INT: ok ok .

INT: cause you look very quiet .

Ms. C1-03: <laugh> .

Ms. C1-03: well [DM] (.) everyone has two sides .

Ms. C1-03: <laugh> .

INT: sure sure .

(Ms. C1-03)

Excerpt 63

It can be seen from both excerpts above that both participants display an inherent stance by using *well* to lead an utterance, respectively. Upon a closer investigation of the context in excerpt 62, the combination of the discourse marker *well* and the utterance "it's his idea", along with another discourse marker *anyway* believed by previous studies (Aijmer, 2004, 2011; Fraser, 2011; Levinson, 1983), accounts for the interpretation of her stance at the moment that she has a repulsion, an abomination, or an objection towards his father's request, asking her to study economics so that it is easier to find a job. This stance can also be encoded from the fact that the participant does not choose the economics as her major in university in the end, as denoted as the consequence of the stance-taking proposed by Du Bois (2007).

Unlike the situation of excerpt 62, it is interesting to notice that the function of *well* in excerpt 63 can be interpreted as a combination of marking an opinion and marking a stance at the same time. The stance shown in this except is attributable to the divergence of the contexts before and after the discourse marker *well*, concerning the opinion of the participant, namely, the saying "everyone has two sides", to the preceding context narrated by the previous speaker in which he is surprised to know that the participant attends a dance activity in contrast to her quiet appearance. Hence,

the participant manifests a stance that it is refutable to make a judgment from people's appearance as everyone has two sides.

Indicating an incomplete answer

This function follows its name, as demonstrated by the previous research (Blakemore, 2014; Müller, 2005), indicating the phenomenon that the utterance preceded by the discourse marker *well* is regarded as an incomplete answer to a prior question. To illustrate this function more explicitly, three key conceptual principles of identifying this function must be noted based on the findings of the previous research (Blakemore, 2014; Fung & Carter, 2007; Müller, 2005; Norrick, 2001). First, a question must be raised in the preceding context before the discourse marker *well*. Second, the current speaker partly answers the question in aligned with the prior question. Third, there will be more explanation after this incomplete answer. Only satisfying all the three principles above by using the discourse marker *well* shall be treated as performing this function. The excerpts below illustrate this function.

INT: then (.) why did you change to the pharmaceutical ?

Ms. C1-03: well [DM] (.) it's because of my parents .

Ms. C1-03: they (.) &-emm (.) ok (.) first (.) they thought it's about medicine .

(Ms. C1-03)

Excerpt 64

INT: but (.) it's (.) just in school .

INT: your life is about (.) study (.) and that's it right ?

Ms. C1-13: well [DM] yeah .

Ms. C1-13: <laugh> .

Ms. C1-13: yeah I just (.) mostly study also .

Ms. C1-13: &-emm (.) but I also did other things too .

(Ms. C1-13)

Excerpt 65

As the two excerpts above show, both utterances preceded by *well* attempt to answer a prior question raised by the previous speaker, respectively. However, it is discernible that the statement "it's because of my parents" in excerpt 64 and a single-word utterance "yeah" in excerpt 65 cannot fulfill the complete answer to the prior questions. In addition, both utterances stated above are followed by more explanation. It must be noted here that due to the lengths of both supplementary explanation in original conversations, the two excerpts only show a part of it, respectively. There is actually more supplementary explanations after the incomplete answers. It is thereby comprehensible that the discourse marker *well* indicates an incomplete answer in this condition.

Indicating an unexpected answer

As the name of this function illustrates, it was discovered that the speaker performed *well* to indicate that the answer to a prior question may be out of expectation from the hearer's perspective (Müller, 2005). Unlike the concept of an incomplete answer analyzed above, an unexpected answer is rendered by the semantic discrepancy between the answer and the question, resulting in the fact that the answer cannot be interpreted as the direct response to the prior question. It is illustrated by the excerpts below.

Ms. C1-03: life is easy here .

Ms. C1-03: I mean [DM] if you like shopping .

Ms. C1-03: but other things (..) like transportation and stuff .

Ms. C1-03: (..) not sure .

INT: what about the people ?

Ms. C1-03: &-emm (..) well [DM] (.) it depends on different people you meet .

INT: oh ok .

(Ms. C1-03)

Excerpt 66

INT: then (.) what kind of teacher should be called (.) a good teacher ?

Mr. C1-22: oh (.) &-emm (…) .

Mr. C1-22: big question .

Mr. C1-22: <laugh> .

INT: yeah (.) think about it .

Mr. C1-22: &-emm (…) well [DM] (.) I'm not sure everyone [/] every student +/ .

Mr. C1-22: +/ has their own sense .

Mr. C1-22: a type of their own (.) they like ?

(Mr. C1-22)

Excerpt 67

In conversational interaction, it is natural for the current speaker to respond to a prior question raised by the previous speaker directly and explicitly. For example, the question raised in excerpt 66 can be interpreted as "what do you think of the people living in Bangkok". It is thus expected by the hearer who proposes this question that the current speaker replies with a description of the people living in Bangkok. On the other hand, it is appropriate to respond to the question raised in excerpt 67 with a description of the quality of being a good teacher. In contrast to answering both

204

questions directly and explicitly, both participants provide an answer whose semantic meaning cannot be consistent with the expected one, respectively. Both utterances preceded by the discourse marker *well* in two excerpts, namely the statement "it depends on different people you meet" in excerpt 66 and "I'm not sure every student has their own sense" in excerpt 67, overtly do not meet the expectation of the previous speakers who raise the questions. It is reasonable to assume, upon the understanding of the semantic meaning of the two utterances, that the participants in both contexts find it difficult or complicated to provide a direct and explicit answer to the prior question. As such, they choose to propose an answer in a way that is out of expectation.

Marking a dispreferred response

It is called a dispreferred response because the utterance addressed by the current speaker embodies a disagreement to a prior utterance stated by the previous speaker (Müller, 2005). It is hence conceivable that the use of the discourse marker *well*, under this circumstance, marks the disagreement in the utterance and mitigates the potential intense between the speaker and the hearer. The excerpt below illustrates this function.

> Ms. C1-03: (.) they are all very rich .
>
> Ms. C1-03: (.) well [DM] (.) not all (.) .
>
> Ms. C1-03: most of them yeah .
>
> Ms. C1-03: just my feeling .
>
> (Ms. C1-03)

Excerpt 68

As demonstrated by the excerpt above, it is relatively simple to identify this function in that the utterance preceded by the discourse marker *well*, namely "not all",

205

manifests a disagreement to the prior utterance "they are all very rich", resulting in two divergent opinions. It is also assumed that the speaker employs *well* to mitigate the following disagreement so that it can help to avoid a face-to-face argument in potential.

In summary, it was delighted to discover that Thai C1-level EFL learners performed the discourse marker *well* in conversation, even though the frequency was low and not many functions were used. Moreover, its interpersonal functions were employed more often than its textual functions, in which marking a stance, a feeling or an opinion was the dominant function.

4.4.2.4 Pragmatic functions of *you know* as a discourse marker

Based on the data of CTELSEC, it was discovered that Thai C1-level EFL learners used one textual function and four interpersonal functions of the discourse marker *you know*, as is summarized in Table 40 and Table 41 below.

Table 40: Textual functions of you know as a discourse marker used by Thai C1-level EFL learners

Textual functions of *you know* as a discourse marker	Number of cases in the raw data
More explanation	4

Table 41: Interpersonal functions of you know as a discourse marker used by Thai C1-level EFL learners

Interpersonal functions of *you know* as a discourse marker	Number of cases in the raw data
To acquire an understanding/an involvement from the hearer	32
An implicature of shared or common knowledge	20

206

An implicature of personal opinion/feeling/stance	8
Leading to an imaginary scene	2

As illustrated by the two tables above, Thai C1-level EFL learners mainly performed its interpersonal functions, wherein two functions, namely, to acquire an understanding or an involvement from the hearer, and an implicature of shared or common knowledge, had the dominant instances.

The details of the identification of each function are presented as follows.

More explanation

It is found that by inserting the discourse marker *you know* intentionally, the speaker provides more information as an explanation to a prior utterance. Hence, there should be an explanation in the utterance preceded by *you know* in terms of a prior utterance, as illustrated by the excerpts below.

> Ms. C1-02: sometimes like [DM] my teacher speaks a little fast .
>
> Ms. C1-02: or (.) sometimes very (..) +/ .
>
> Ms. C1-02: +/ very difficult .
>
> Ms. C1-02: and I can just ask my friends who sit next to me .
>
> Ms. C1-02: you know [DM] (.) just speak quietly without teachers seeing yeah .
>
> (Ms. C1-02)

Excerpt 69

> Mr. C1-09: people just come here to live their life .
>
> Mr. C1-09: this is their first choice (.) I guess .
>
> Mr. C1-09: you know [DM] (.) they can earn more money here .
>
> Mr. C1-09: many more buildings up .

(Mr. C1-09)

Excerpt 70

As both excerpts shown above, the utterance preceded by the discourse marker *you know* includes an explanation in detail of a prior utterance, respectively. In excerpt 69, the participant explains how she can get help from her friends studying in a classroom, while the participant in excerpt 70 explains why many Thai people come to Bangkok to live. It can be seen that the explanation with details in both contexts is delivered in the lead of a discourse marker *you know*.

An important point must be addressed regarding the identification of this function is that to some extent, the function of the discourse marker *you know* in this type of context can also be interpreted as to acquire an understanding or an involvement from the hearer in that the utterance preceded by *you know* provides new information to intrigue the hearer's attention (House, 2009). This research cannot deny this fact in that the multi-functionality of the discourse marker *you know* results in the multi-interpretations from different researchers' points of view. Under this circumstance, the Relevance Theory must be adopted at this point, proposing a conceptual principle that it is relevant enough for it to be worth the hearer's effort to process it and it is the most relevant one compatible with the speaker's ability and preference (Sperber & Wilson, 1986). It is from the researcher's perspective to examine the utterances within the given context that the speaker overtly wants to provide more explanation to the prior utterance. Hence, the use of the discourse marker *you know* in this context is classified under this type of function.

To acquire an understanding/an involvement from the hearer

The same as the situation of Thai B1-level EFL learners, this function is also performed at the highest frequency by Thai C1-level EFL learners. The excerpt extracted from CTELSEC illustrates this function.

Mr. C1-01: and (.) that's the only thing we were good at .

Mr. C1-01: (.) now (.) we didn't know what other things we are good at .

Mr. C1-01: you know [DM] .

(Mr. C1-01)

Excerpt 71

Upon the investigation of the context, the participant in excerpt 71 addresses the fact that some students can only study, yet they are not good at anything else. The identification of the function of the discourse marker *you know* in this context is thus rendered by the provided information followed by a *you know* at the end of it, reflecting that the participant acquires an understanding from the hearer at this point.

An implicature of shared or common knowledge

It can be seen from the data that Thai C1-level EFL learners performed this function more frequently than Thai B1-level EFL learners. The two excerpts below illustrate this function.

Mr. C1-01: you know [DM] (.) the connection can't always be good .

Mr. C1-01: sometimes it's [/] it's disconnected .

Mr. C1-01: oh right .

(Mr. C1-01)

Excerpt 72

Mr. C1-02: yeah (.) the trees and big waterfall .

Mr. C1-02: they have everything .

Mr. C1-02: you know [DM] (.) their airport is so amazing and beautiful .

(Ms. C1-02)

Excerpt 73

It is overtly to see that the utterances preceded by the discourse marker *you know* in both contexts are in reference to two pieces of common knowledge: the common knowledge that the signal of WIFI in campus cannot always be good in excerpt 72, and Singapore's airport is well-known to the worldwide as it is amazing and beautiful in excerpt 73. The finding is aligned with the previous research (Müller, 2005; Vanda & Peter, 2011) that the shared or common knowledge may connect the interpersonal relationship to all the interlocutors in conversation.

An implicature of personal opinion/feeling/stance

Not many cases were discovered to be in this function, as illustrated by the two excerpts below.

Ms. C1-02: I may need to tell myself (..) it [//] today is another day .

Ms. C1-02: today is going to be fine .

Ms. C1-02: it's another day .

Ms. C1-02: and (.) you know [DM] maybe I can feel better .

(Ms. C1-02)

Excerpt 74

Ms. C1-19: if you don't know how to do it .

Ms. C1-19: you know [DM] (.) I prefer you just tell me .

INT: oh I know your feeling .

(Ms. C1-19)

Excerpt 75

As the two excerpts show above, it is conceivable that the discourse marker *you know* is employed to indicate the personal opinion in excerpt 74 and a stance in excerpt 75, respectively. This function is also illustrated by the previous research the same as the finding shown above (Müller, 2005; Vanda & Peter, 2011).

Leading to an imaginary scene

This function refers to the phenomenon that a series of utterances preceded by the discourse marker *you know* are narrated in reference to a description of a scene where some details are presented (Müller, 2005). It is responsible for various purposes in conversation, i.e., to present some details to a prior utterance; to describe a related topic; to achieve a better understanding, etc. It is interesting to identify two cases in this research that perform this function, as illustrated below.

Ms. C1-02: I sometimes can't always listen to teacher .

Ms. C1-02: you know [DM] in classroom (.) you can have interact .

Ms. C1-02: the classmates talking .

Ms. C1-02: the teacher's movement .

(Ms. C1-02)

Excerpt 76

Ms. C1-02: but of course sometimes we would be out of control .

Ms. C1-02: and (.) you know [DM] there might be some Thai teachers .

Ms. C1-02: they will shout to us .

Ms. C1-02: shut up .

Ms. C1-02: be quiet .

(Ms. C1-02)

Excerpt 77

It is comprehensible from the two given contexts that s series of utterances combines with respect to a scene with certain details given by the participant. The given scene in excerpt 76 refers to studying in a classroom in detailed information like "have interact", "classmates talking" and "teacher's moving", while the provided scene in excerpt 77 regards an imaginary Thai teacher's behavior in a classroom with the detailed description such as "shout to us", "shut up" and "be quiet".

In summary, Thai C1-level EFL learners used the discourse marker *you know* mainly for its interpersonal functions as Thai B1-level EFL learners did.

4.4.2.5 Pragmatic functions of *I think* as a discourse marker

It was found that Thai C1-level EFL learners used two functions of the discourse marker *I think*, both of which are interpersonal functions. Table 42 summarizes the function of *I think* as a discourse marker used by Thai C1-level EFL learners.

Table 42: Pragmatic functions of I think as a discourse marker used by Thai C1-level EFL learners

Pragmatic functions of *I think* as a discourse marker	Number of cases in the raw data
An implicature of personal opinion/feeling/stance	204
A hedge/mitigator	2

As is shown in the table above, like the situation of Thai B1-level EFL learners, an implicature of personal opinion, feeling and stance found its dominant way as the major function used by Thai C1-level EFL learners, while two cases were found to use *I think* as a hedge or a mitigator in conversation.

The details of the identification of each function are presented as follows.

An implicature of personal opinion/feeling/stance

Like the situation of Thai B1-level EFL learners, it was found that Thai C1-level EFL learners had a preference to employ *I think* to express their personal opinion, feeling or stance, as illustrated in the excerpts below.

> Mr. C1-11: &-uhh (.) you mean Bangkok ?
>
> INT: yes Bangkok .
>
> Mr. C1-11: oh &-emm (.) it's crowded .
>
> Mr. C1-11: much more pollution .
>
> Mr. C1-11: I think [DM] (.) all the [/] the situations now are not so good .
>
> (Mr. C1-11)

Excerpt 78

> Ms. C1-19: I don't remember a lot .
>
> INT: oh it's fine .
>
> Ms. C1-19: yeah (.) it's (..) ok (.) fine (.) I think [DM] .
>
> (Ms. C1-19)

Excerpt 79

The two excerpts above show that both participants have the comment clauses next to a discourse marker *I think*, reflecting their opinion or stance to a prior topic discussed in the preceding context. It is noticed that the collocational pattern *I think* is detached from the comment clauses, with the longer pause surrounding it, the changeable position of either at the start or at the end of an utterance and *that*-deletion. It is therefore conceivable that *I think* serves as a discourse marker in this condition.

A hedge/mitigator

It was found that *I think* was used as a hedge or a mitigator in the context where the speaker triggered a disagreement or remains a contrary statement in which a potential argument may appear at the moment (Dehe & Wichmann, 2010; Diessel & Tomasello, 2001). It is thus assumed that the speaker wants to mitigate the atmosphere. Hence, it should be noted concerning this function that a disagreement or a contrary opinion must be preceded by the discourse marker *I think*, as illustrated below.

INT: you need to think about if you make enough money to support yourself .

INT: is it a right idea .

Ms. C1-24: but I think [DM] Singapore can have a better (.) higher salary .

Ms. C1-24: isn't it ?

Ms. C1-24: I can support myself .

INT: oh ok .

(Ms. C1-24)

Excerpt 80

Ms. C1-28: yeah (.) really beautiful .

INT: but I'm not sure they photoshop it .

Ms. C1-28: &-uhh (.) I think [DM] it's real .

Ms. C1-28: I don't see they are photoshopped .

Ms. C1-28: I can see it as natural as that .

Ms. C1-28: it should be .

INT: ok ok .

(Ms. C1-28)

Excerpt 81

Upon a detailed examination of the contexts in both excerpts above, it can be seen that a contrary opinion is illustrated preceded by *I think*, respectively. In excerpt 80, the previous speaker tries to persuade the participant to consider working in Singapore in the preceding context, whereas the participant states a contrary opinion by using a combination of *but* and *I think*, meaning that she considers working in Singapore has no problem. In excerpt 81, the previous speaker has doubts on a seemingly photo-shopped picture, whereas the participant denies this thought by stating the utterance "it's real" as a contrary opinion preceded by a filler *uhh* and *I think* with some longer pauses surrounding them. Hence, it can be interpreted that both participants intend to ease off the potential intense from the contrary opinion by the fact that they insert the discourse marker *I think*, with the conjunction of contradiction *but* or a filler *uhh*, respectively.

In summary, Thai C1-level EFL learners used the discourse marker *I think* as an implicature of personal opinion, feeling or stance in its predominance.

4.4.2.6 Pragmatic functions of *I mean* as a discourse marker

The data of CTELSEC shows that Thai C1-level EFL learners used four textual functions and three interpersonal functions of the discourse marker *I mean*, respectively, as summarized in Table 43 and Table 44 below.

Table 43: Textual functions of I mean as a discourse marker used by Thai C1-level EFL learners

Textual functions of *I mean* as a discourse marker	Number of cases in the raw data
More explanation/exemplification	40
Marking a repair	25
Marking a topic shift	5

Searching for the right word	1

Table 44: Interpersonal functions of I mean as a discourse marker used by Thai C1-level EFL learners

Interpersonal functions of *I mean* as a discourse marker	Number of cases in the raw data
An implicature of personal opinion/feeling/stance	27
An implicature of the cause subjectively	15
Emphasis/repetition	3

It is observed from the two tables above that the textual function, namely, marking more explanation or exemplification, was performed as the dominant function. In the meantime, marking a repair and an implicature of personal opinion, feeling or stance were also performed in a high frequency. Four new functions were performed by Thai C1-level EFL learners compared to the use of it by Thai B1-level EFL learners.

The identification of each function is presented as follows.

More explanation/exemplification

As noted earlier, the utterance preceded by the discourse marker *I mean* provides more explanation or exemplification to a prior utterance in the preceding context, as illustrated by the two excerpts as follows.

Mr. C1-06: I kind of cheated a lot .

Mr. C1-06: but not on the exam (.) though .

Mr. C1-06: I mean [DM] (..) how (.) I lived my life as a student .

Mr. C1-06: for example the attendance of the classes .

Mr. C1-06: or (..) yeah (.) other things .

(Mr. C1-06)

Excerpt 82

Mr. C1-27: they just use what they have .

Mr. C1-27: that's quite sad right ?

Mr. C1-27: I mean [DM] from my experience .

(Mr. C1-27)

Excerpt 83

In excerpt 82, the participant provides more explanation by the utterance "how I lived my life as a student" followed by an example to explain the preceding context where the participant utters the topic of cheating behavior as a student. In excerpt 83, the participant utters an explanation to the preceding context, stating the utterance "from my experience".

It is worth mentioning that it is assumed that this function is performed dominantly in that it is derived from the semantic meaning of the NP form *I mean*, as also cited by the previous research (Brinton, 2010).

Marking a repair

Upon the examination of the data of CTELSEC, Thai C1-level EFL learners perform this function the same as Thai B1-level EFL learners, as illustrated by the excerpts below.

Mr. C1-01: don't (.) focus too much on what you study .

INT: is that right ?

Mr. C1-01: yeah (.) OK sorry.

Mr. C1-01: I mean [DM] (.) don't be a nerd too much .

(Mr. C1-01)

Excerpt 84

Mr. C1-06: (.) and [/] and become a teacher .

Mr. C1-06: I mean [DM] not a teacher .

Mr. C1-06: a professor in university .

(Mr. C1-06)

Excerpt 85

It is interesting to observe in excerpt 84 that the participant apologizes for his error in the prior utterance before having a repair of it preceded by a discourse marker *I mean*, reflecting that he has an intention to have this repair. Meanwhile, it is also simple to identify the function of *I mean* in excerpt 85 because the participant clearly states the error in the prior utterance and repairs it by "a professor in university".

Marking a topic shift

This function is performed at the moment when the speaker shifts the previous topic into a new one (Brinton, 2010). Two excerpts below illustrate this function.

Ms. C1-15: I don't really need to go abroad .

Ms. C1-15: I mean [DM] (.) I have some friends .

Ms. C1-15: they (.) have the purpose to go abroad .

Ms. C1-15: they (.) like that .

Ms. C1-15: it's their choose .

(Ms. C1-15)

Excerpt 86

Mr. C1-27: but I love this country .

Mr. C1-27: cause it's beautiful .

Mr. C1-27: life here is easy (.) I guess .

Mr. C1-27: I mean [DM] if it's Bangkok .

Mr. C1-27: then probably not so easy .

Mr. C1-27: people are suffering (.) and (.) working hard .

(Mr. C1-27)

Excerpt 87

In excerpt 86, the participant illustrates that going abroad is not essential for her. She then shifts the topic into the introduction of her friends' choices by stating a series of utterances preceded by the discourse marker *I mean*. In excerpt 87, the participant firstly discuss the lifestyle living in the country in the preceding utterance before the discourse marker *I mean*. The topic is shifted after *I mean*, illustrating the living situation of Bangkok.

A point worth noting here is that for the topic shift led by the discourse marker *I mean*, it can be seen that the topics before and after it are relevant to some extent, i.e., it is about the choice of going abroad in excerpt 86 and about the living condition in excerpt 87. However, concerning the shift of the subject, the event and the scene, as demonstrated by the previous research (Erman, 1987; Fox et al., 2002; Vanda & Peter, 2011), it is thus regarded as marking a topic shift.

Searching for the right word

Only one case was found in CTELSEC where this function was performed, as shown in the excerpt below.

Ms. C1-15: I just (.) wish she could change her mind .

Ms. C1-15: maybe just a little .

Ms. C1-15: I mean [DM] (..) &-uhh (.) what can I say .

Ms. C1-15: maybe she could be more modern (.) I guess .

219

(Ms. C1-15)

Excerpt 88

It is interesting to observe that to search for the right word, the participant in the excerpt above employs several strategies, including two discourse markers *I mean* and *I guess* inserted at the beginning and at the end of the utterance "maybe she could be more modern", respectively, with longer pauses, a filler *uhh* and an utterance "what can I say" to reflect her on-going search at the moment evidently.

An implicature of personal opinion/feeling/stance

Like the other discourse markers under investigation in this research, such as *you know*, *I mean* as a discourse marker can also be used to indicate personal opinion, feeling or stance, as illustrated by the excerpt below.

> Mr. C1-06: we need to share (.) one floor .
>
> INT: oh (.) is it convenient ?
>
> Mr. C1-06: it's (.) ok .
>
> Mr. C1-06: I mean [DM] I don't think it's a problem .
>
> Mr. C1-06: but just for me .
>
> (Mr. C1-06)

Excerpt 89

In excerpt 89, the discourse marker *I mean* is performed followed by the expression *I don't think* to indicate that the participant does not consider that using a shared bathroom is a problem, denoting his stance to the issue in the preceding context.

An implicature of the cause subjectively

It was argued that the discourse marker *I mean* may lead some causes for a prior utterance, wherein the reasons are given from the speaker's perspective (Brinton, 2010; Vanda & Peter, 2011). To identify this function, it is assumed that a cause-and-effect relationship must be adopted, as illustrated by the two excerpts extracted from the data as follows.

Mr. C1-01: but now that I look back .

Mr. C1-01: it [//] it's (.) just not so great .

Mr. C1-01: &-emm (.) I mean [DM] (.) my friends my teachers .

Mr. C1-01: and (.) yeah (.) the atmosphere are not so good in overall .

(Mr. C1-01)

Excerpt 90

Ms. C1-15: I can't live like that .

Ms. C1-15: I mean [DM] I need to have supermarket around .

Ms. C1-15: so that I can but something if I want to have something .

(Ms. C1-15)

Excerpt 91

In both excerpts shown above, both participants present the reasons for the prior utterance from their own perspectives, respectively. The participant in excerpt 90 does not think his previous university life is good because some of his friends, teachers and the atmosphere around him are not good. The participant in excerpt 91 cannot live in the western style because she needs supermarkets around to buy things at any time rather than driving miles to look for groceries. It can be interpreted as the cause-and-effect relationship between the utterance preceded by the discourse marker *I mean* and the utterance followed by it.

221

Emphasis/repetition

It was found that the discourse marker *I mean* was performed to emphasize or mark a repetition of some information (Brinton, 2010). It is relatively easy to identify in that the information should be emphasized by the tone or the reiteration of it, as illustrated by the excerpt below.

Mr. C1-01: but that was me .

Mr. C1-01: &-uhmm (..) I [/] I know I feel I have .

Mr. C1-01: &-uhmm (.) friends who are very rich .

INT: ok [>] .

Mr. C1-01: [<] back then .

Mr. C1-01: and I mean [DM] (.) they are rich rich .

(Mr. C1-01)

Excerpt 92

As demonstrated by the excerpt above, it is observed that the word "rich" is repeated several times in the context, reflecting the emphasis of this information by the participant. The reiteration of it in the utterance preceded by *I mean* denotes that the participant intends to emphasize this information to the hearer.

To summarize, Thai C1-level EFL learners used the discourse marker *I mean* in seven functions, in which its textual functions were more frequently used than its interpersonal functions. Three functions, namely, marking more explanation or exemplification, an implicature of personal opinion, feeling or stance, and marking a repair, were employed the most, whereas other functions were used in lower frequencies.

4.4.3 A conclusion of the pragmatic functions used by Thai EFL learners

In summary, as the analysis above presents, a total of 39 pragmatic functions were identified in all the six English spoken discourse markers investigated in this research, wherein 21 textual functions and 18 interpersonal functions were used by Thai EFL learners, respectively.

For Thai B1-level EFL learners, 15 textual functions and 9 interpersonal functions were identified, while for Thai C1-level EFL learners, 19 textual functions and 17 interpersonal functions were identified. Therefore, it is generally seen that Thai C1-level EFL learners performed more pragmatic functions of the English spoken discourse markers than Thai B1-level EFL learners.

It should be noted, as illustrated in the analysis of the pragmatic functions of each spoken discourse marker, that 9 out of 39 pragmatic functions were used fewer than three instances, i.e., the function of *so*: more explanation with one instance; three functions of *you know*, searching for the right words with one instance, emphasis or repetition with one instance, leading a summary with two instances; two functions of *well*, searching for the right words with two instances, marking a dispreferred response with three instances; the function of *I think*, as a mitigator with two instances; two functions of *I mean*, searching for the right words with one instance, emphasis or repetition with three instances. Considering the issue of generalization in the use of spoken discourse markers, it can be concluded that 30 out of 39 pragmatic functions were generally used by Thai EFL learners, while 9 pragmatic functions were rarely used.

After the analysis of the pragmatic functions of each spoken discourse markers used by Thai EFL learners, in the next part, question 4, namely, what are the different pragmatic functions of spoken discourse markers used by Thai EFL learners compared to native English speakers, will be answered.

223

4.5 Question 4: What are the different pragmatic functions of spoken discourse markers used by Thai EFL learners compared to native English speakers?

The previous part identified the pragmatic functions of each spoken discourse marker investigated in this research used by Thai EFL learners. Regarding the fourth research question, the research compared the pragmatic functions of each spoken discourse marker used in both TELSEC and SBCSAE.

This part presents the differences in the use of each spoken discourse marker between TELSEC and SBCSAE, and has a conclusion of all the major differences in using the pragmatic functions of spoken discourse markers between Thai EFL learners and native English speakers from three facets as follows:

First, this research provides the LL test results of both overall functions and each function of each spoken discourse marker to examine what functions have a significant difference between both corpora.

Next, this research illustrates the raw frequency and the corresponding proportion of each function of each spoken discourse marker, and the normalized frequency of each function of each spoken discourse marker used by both Thai EFL learners and native English speakers to examine what functions are used differently between both corpora.

Finally, this research compares the number and the corresponding proportion of the participants in both corpora who used the same function of each spoken discourse marker. This is because if the number or proportion of Thai EFL learners who used a function is little, it may also reflect the deficiency in using this function of the spoken discourse marker.

From the three facets illustrated above, it is believed that the differences in using the spoken discourse markers between both corpora can be seen comprehensively.

In the following parts, comparisons of the three facets in each spoken discourse marker will be presented one after another.

4.5.1 Different pragmatic functions of *like* as a discourse marker

Table 45 illustrates the LL test results of each function and all the functions of *like*, the raw frequency and its corresponding proportion of each function and all used functions of *like*, the normalized frequency of each function and all used functions of *like*, and the number and its corresponding proportion of the participants who used each function and all the functions of *like* in both TELSEC and SBCSAE.

It is worth mentioning that the order of the functions in the table is based on the highest raw frequency to the lowest raw frequency of each function used by native English speakers in SBCSAE. This is because SBCSAE was used as the native English speaker corpus to be the reference for comparing the learner corpus built in this research. It is therefore considered that the way of using each spoken discourse marker should be the way in which native English speakers used in SBCSAE. All the tables presented in this section follow the principle of this order.

Meanwhile, in each table, TF stands for textual function. IF stands for interpersonal function. RF stands for raw frequency. P stands for proportion. NP stands for normalized frequency. NP stands for the number of participants.

Table 45:Tabl Comparisons of each function and all functions of like between SBCSAE and TELSEC

like	SBCSAE			TELSEC			LL	SBCSAE		TELSEC	
Function (all are TF)	RF	P (%)	NF	RF	P (%)	NF		NP	P %	NP	P %
To exemplify	388	26.4	16	115	26.8	15	- 0.03	182	61.1	16	26.7
Searching for the right words	380	25.9	15	15	3.5	2	- 116.77	205	68.8	4	6.7
Marking a focus on new information	317	21.6	13	187	43.6	25	+ 49.12	132	44.3	25	41.7
As an approximator	244	16.6	10	66	15.4	9	- 0.64	121	40.7	21	35.0
To have a quotation	100	6.8	4	26	6.1	3	- 0.47	71	23.8	14	23.3
More information/explanation	32	2.2	1	20	4.7	3	+ 6.06	17	5.7	9	15
A repair	8	0.5	0	0	0	0	- 18.46	7	2.3	0	0
Total	1469	100	59	429	100	57	- 0.36	229	76.8	28	46.7
Total number of participants in the corpus								298	100	60	100

Figure 9 illustrates the proportion of each function of *like* used by Thai EFL learners and native English speakers, while Figure 10 illustrates the proportion of the participants who each function of *like* used by Thai EFL learners and native English speakers.

Figure 9: Proportion of each function of the discourse marker like used by Thai EFL learners and native English speakers
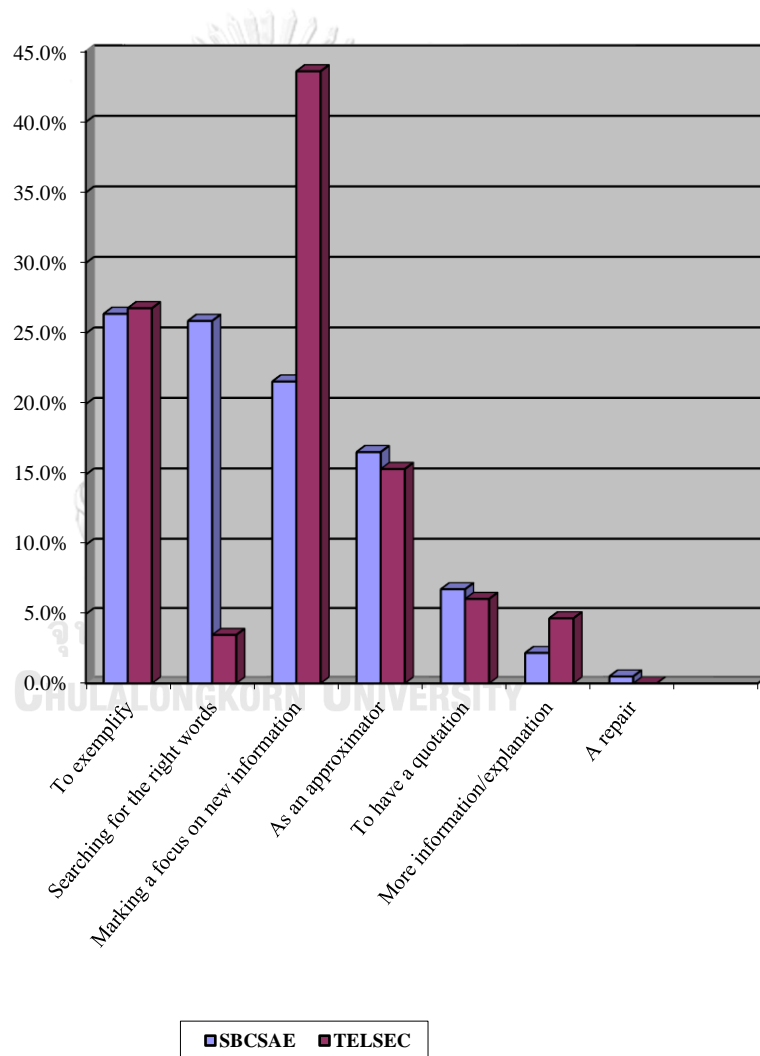
Figure 10: Proportion of participants who used the function of the discourse marker
like in TELSEC and SBCSAE
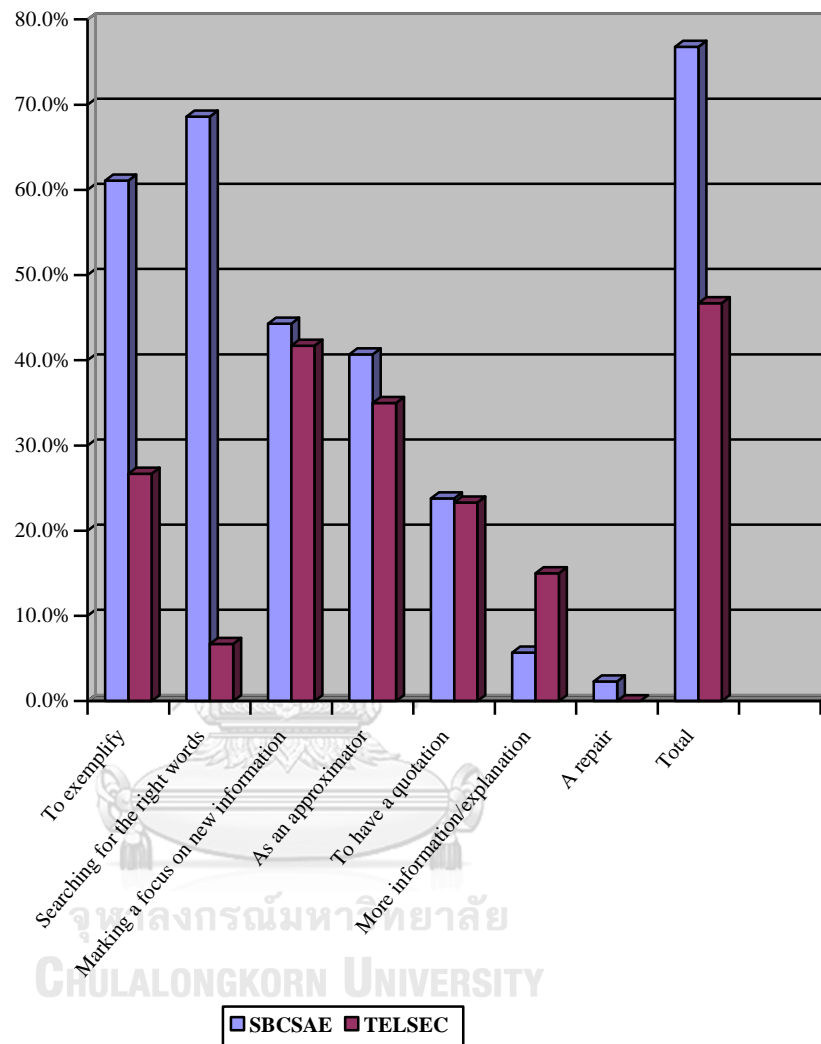
Based on the table and figures shown above, a huge difference was found in the function: searching for the following words, wherein a limited number of Thai EFL learners used it in a low frequency while a large amount of native English speakers used it at a much higher frequency. Meanwhile, the proportion of Thai EFL learners who used the function, to exemplify, was less than half of native English speakers

even though there was no significant difference in using this function between the two corpora. In contrast, Thai EFL learners focused on the function: marking a focus on new information, with the highest frequency by the largest proportion of participants. On the contrary, native English speakers also used it in a high frequency, but not so high as Thai EFL learners. As shown by the data, the function: as a repair, was excluded from TELSEC. Meanwhile, based on the previous research (Diskin, 2017), the interpersonal function: as a mitigator, was not discovered in both corpora.

Based on the data, it is interesting to discover that the function of *like* used by Thai EFL learners in the highest frequency, namely, marking a focus on new information, was different from the one used by native English speakers, namely, to exemplify. Since the proportion of the participants who used the function, namely, marking a focus on new information, was almost the same in both corpora, it is suggested that Thai EFL learners tended to stick to using this function in conversation. Meanwhile, even though there was no significant difference in using the function, to exemplify, the proportion of Thai participants was more than half of native English speakers, which suggests that this function was only used by a small amount of Thai EFL learners. Moreover, the largest difference was the use of function, searching for the right words, with a significant difference and only four participants' use. This result suggests that Thai EFL learners lack the knowledge of this function of *like* in conversation. Instead of using the fillers, such as *ehh*, *uhmm*, etc., the spoken discourse marker *like* helps EFL learners as a signal that they are searching for the next words at the moment in conversation. Lastly, based on no use of the function, as a repair, Thai EFL learners may need guidance for understanding this function.

As similar as the result of Diskin (2017), for non-native English speakers who used the spoken discourse marker *like*, the function, to exemplify, was used in a high frequency. However, it is interesting to see that the function, more information or

229

explanation, was used in the highest frequency in Diskin (2017), while was relatively used in a low frequency by Thai EFL learners. This dissimilarity may suggest the variety of using spoken discourse markers by different groups of non-native English speakers. Meanwhile, Diskin (2017) discovered one interpersonal function of *like*, namely, as a mitigator, which was not found in both SBCSAE and TELSEC. Based on the analysis of Diskin (2017), this interpersonal function was mostly used by Irish in Irish English, which was considered the use of this function in a regional limitation. This may explain the reason why this interpersonal function was not found in both corpora in this research. Hence, this research will not discuss this function deeply.

4.5.2 Different pragmatic functions of *so* as a discourse marker

Table 46 illustrates the LL test results of each function and all the functions of *so*, the raw frequency and its corresponding proportion of each function and all used functions of *so*, the normalized frequency of each function and all used functions of *so*, and the number and its corresponding proportion of the participants who used each function and all the functions of *so* in both TELSEC and SBCSAE.

Table 46: Comparisons of each function and all functions of so in SBCSAE and TELSEC

so	SBCSAE			TELSEC			LL	SBCSAE		TELSEC	
Function	**RF**	**P (%)**	**NP**	**RF**	**P (%)**	**NP**	**LL**	**NP**	**P (%)**	**NP**	**P (%)**
Marking a question (IF)	370	21.2	15	3	1.6	0	- 169.05	217	72.8	3	5.0
Marking a result/consequence (TF)	307	17.6	12	68	36.1	9	- 5.68	170	57.0	11	18.3
Marking opinion/feeling/stance (IF)	274	15.7	11	39	20.7	5	- 23.18	146	49.0	7	11.7
Leading a summary (TF)	221	12.7	9	19	10.1	3	- 39.30	194	65.1	6	10.0
A transition sequentially (TF)	162	9.3	7	11	5.9	1	- 35.71	74	24.8	2	3.3
Prompting the hearer to speak/an implied result (IF)	122	7.0	5	6	3.2	1	- 33.46	122	40.9	3	5.0
Hold the floor (IF)	106	6.1	4	6	3.2	1	- 26.67	98	32.9	2	3.3
Marking a start of a narration (TF)	67	3.8	3	30	16.0	4	+ 3.05	65	21.8	16	26.7
Marking a new topic (TF)	63	3.6	3	0	0	0	- 33.23	57	19.1	0	0
Marking a self-correction/a repair (TF)	18	1.0	1	0	0	0	- 9.50	7	2.3	0	0
More explanation (TF)	17	1.0	1	1	0.5	0	- 4.17	16	5.4	1	1.7
Topic shift back (TF)	16	0.9	1	5	2.7	1	+ 0.00	8	2.7	3	5.0
Total	1743	100	70	188	100	25	- 236.18	231	77.5	24	40.0
Total number of participants in the corpus								298	100	60	100

Figure 11 illustrates the proportion of each function of *so* used by Thai EFL learners and native English speakers, while Figure 12 illustrates the proportion of the participants who used each function of *so* used by Thai EFL learners and native English speakers.

Figure 11: Proportion of each function of the discourse marker so used by Thai EFL learners and native English speakers
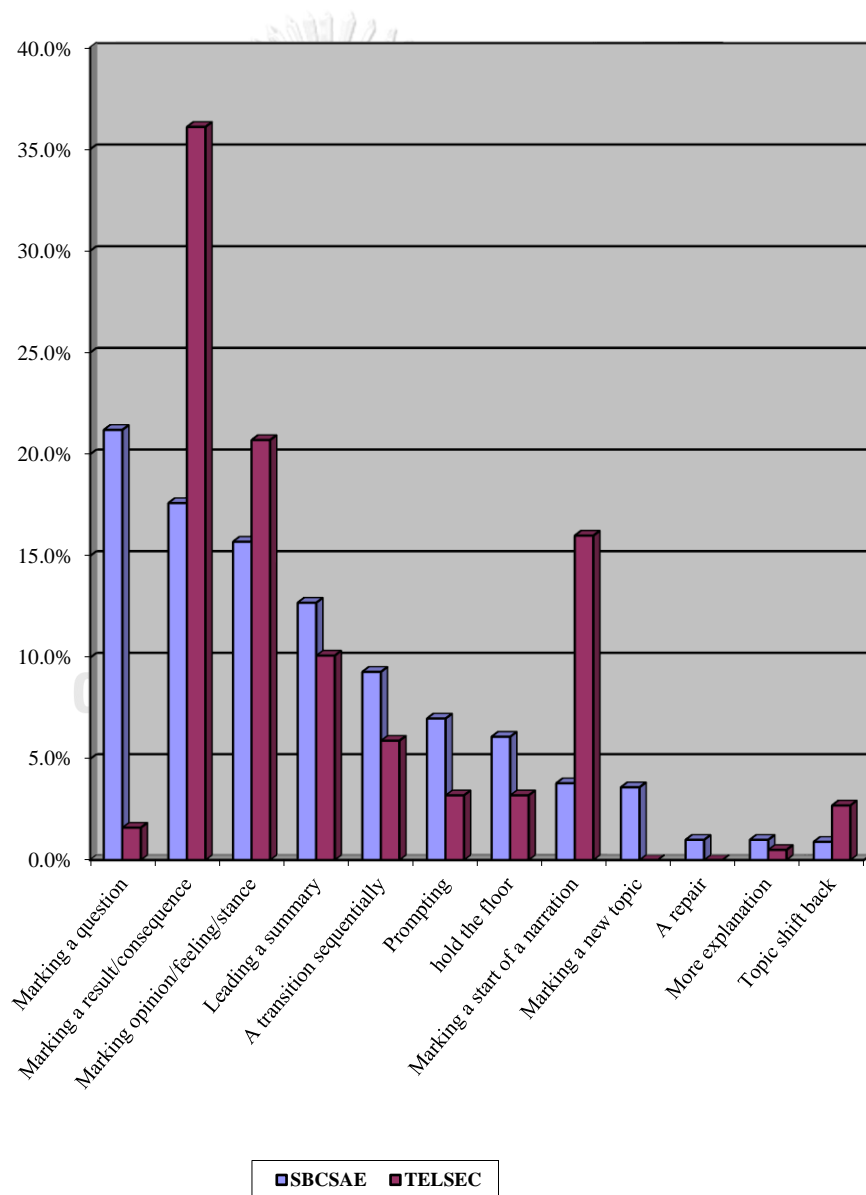
Figure 12: Proportion of participants who used the function of the discourse marker
so in TELSEC and SBCSAE
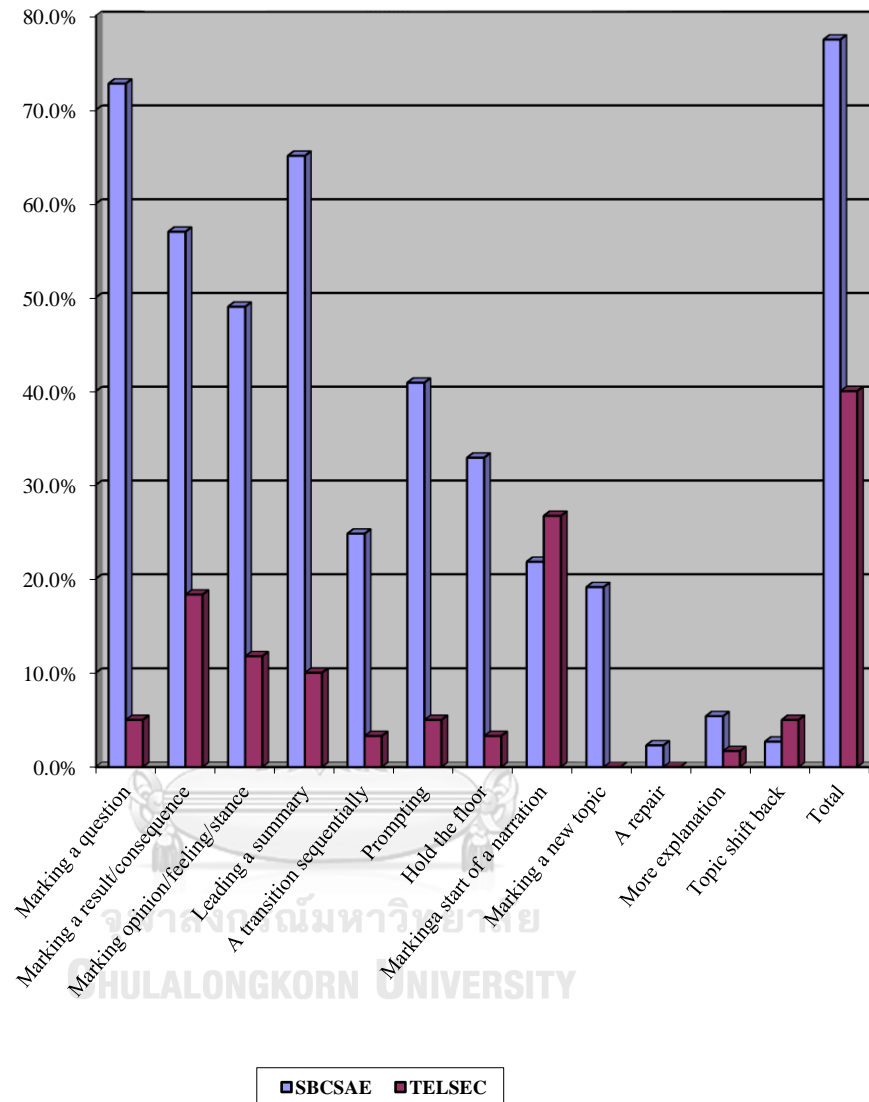
Based on the table and figures shown above, it is observed that the biggest difference was in the function, marking a question, with only three Thai EFL learners using it in a low frequency compared to native English speakers. On the contrary, Thai EFL learners focused on the two functions, namely, marking a result or

233

consequence and marking a start of a new narration, more than native English speakers. As is shown in the table, the proportion of Thai participants who used the function, marking a result or consequence, was only one-third of native English speakers, while the raw frequency was twice higher than native English speakers. It indicates that only a limited range of Thai EFL learners used this function, and each Thai participant who used this function averagely tended to stick to using it. Meanwhile, 10 out of 12 functions shared the phenomenon that the proportions of the native English speakers who used them were twice to seven times higher than the proportions of Thai EFL learners, wherein five of them were textual functions and the other five were interpersonal functions. It reflects that only a limited number of Thai EFL learners used *so* as a discourse marker in communication.

As illustrated in previous research (Buysse, 2012), it was believed that the function, marking a result or consequence, was one of the commonly used functions of *so*, which was aligned with the result of this research. However, as this research deeply analyzed the number of participants who used each function in both corpora, it can be seen that the number of Thai EFL learners who used this function was much fewer than native English speakers. Based on Table 45 and Figure 12, 83% of the functions of *so* (10 out of 12 functions) shared a huge difference in the proportion of participants who used each function. This leads to the difference in the total participants who used *so* in which the proportion of Thai participants who used *so* was half of native English speakers. Besides, based on the data in this research and the result of previous research (Buysse, 2012; Müller, 2005), two interpersonal functions, namely, marking a question and prompting the hearer to speak, were used at a high frequency in both native English speakers and non-native English speakers. It was found, in this research, that both interpersonal functions were used at a low frequency by Thai EFL learners. It may indicate that Thai EFL learners lack the

ability to use interpersonal functions of spoken discourse markers in conversation to connect the relationship with the hearer.

4.5.3 Different pragmatic functions of *well* as a discourse marker

Table 47 illustrates the LL test results of each function and all the functions of *well*, the raw frequency and its corresponding proportion of each function and all used functions of *well*, the normalized frequency of each function and all used functions of *well*, and the number and its corresponding proportion of the participants who used each function and all the functions of *well* in both TELSEC and SBCSAE.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Table 47: Comparisons of each function and all functions of well in SBCSAE and TELSEC

well Function	SBCSAE			TELSEC			LL	SBCSAE		TELSEC	
	RF	P (%)	NP	RF	P (%)	NP		NP	P (%)	NP	P (%)
All the other interpersonal functions (IF)	461	33.1	19	0	0	0	- 243.19	177	59.4	0	0
Marking opinion/feeling/stance (IF)	252	18.1	10	7	29.2	1	- 89.04	201	67.4	3	5.0
Searching for the right words (TF)	214	15.3	9	2	8.3	0	- 95.36	133	44.6	1	1.7
Marking a dispreferred answer (IF)	135	9.7	5	3	12.5	0	- 51.08	123	41.3	2	3.3
All the other textual functions (TF)	116	8.3	5	0	0	0	- 41.96	100	33.6	0	0
Marking an unexpected answer (IF)	91	6.5	4	4	16.7	1	- 26.53	88	29.5	2	3.3
Marking an incomplete answer (IF)	83	6.0	3	4	16.7	1	- 23.03	82	27.5	2	3.3
Marking a new topic (TF)	42	3.0	2	4	16.7	1	- 6.67	42	14.1	2	3.3
Total	1394	100	56	24	100	3	- 562.17	214	71.8	3	5.0
Total number of participants in the corpus								298	100	60	100

236

Figure 13 illustrates the proportion of each function of *well* used by Thai EFL learners and native English speakers, while Figure 14 illustrates the proportion of the participants who used each function of *well* used by Thai EFL learners and native English speakers.

Figure 13: Proportion of each function of the discourse marker well used by Thai EFL learners and native English speakers
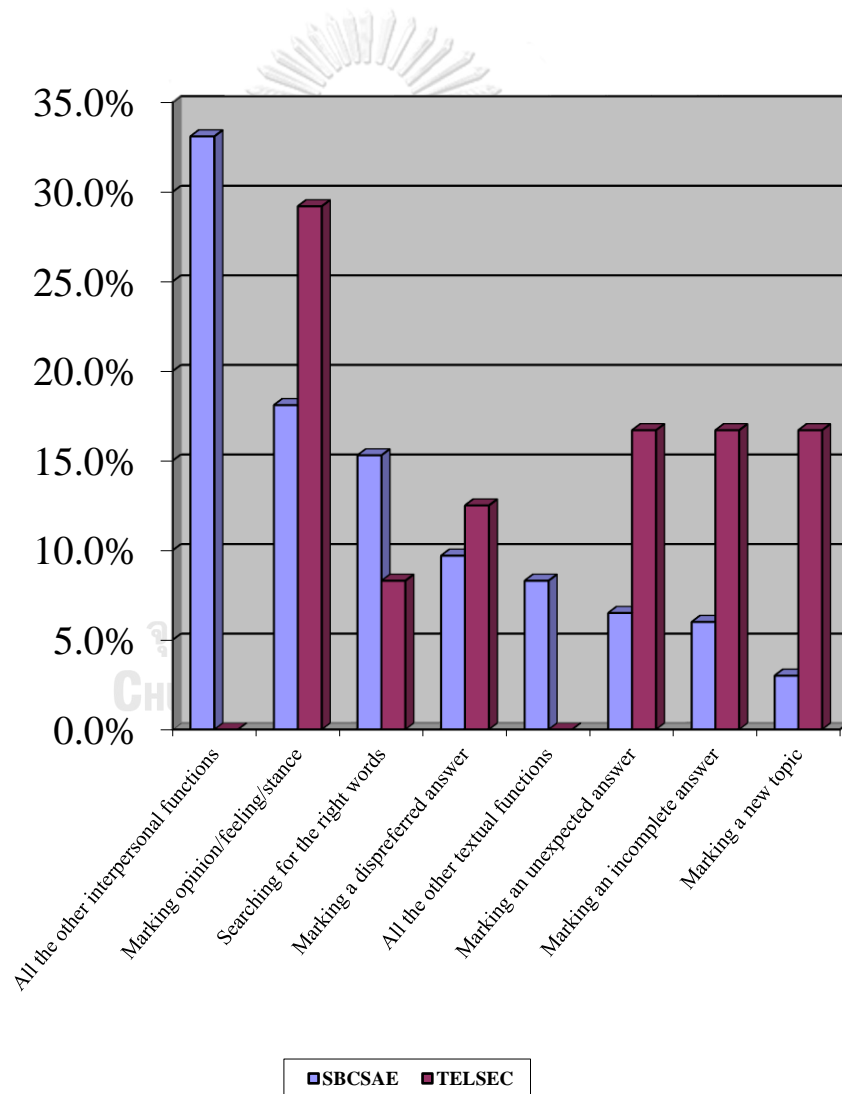
Figure 14: Proportion of participants who used the function of the discourse marker well in TELSEC and SBCSAE
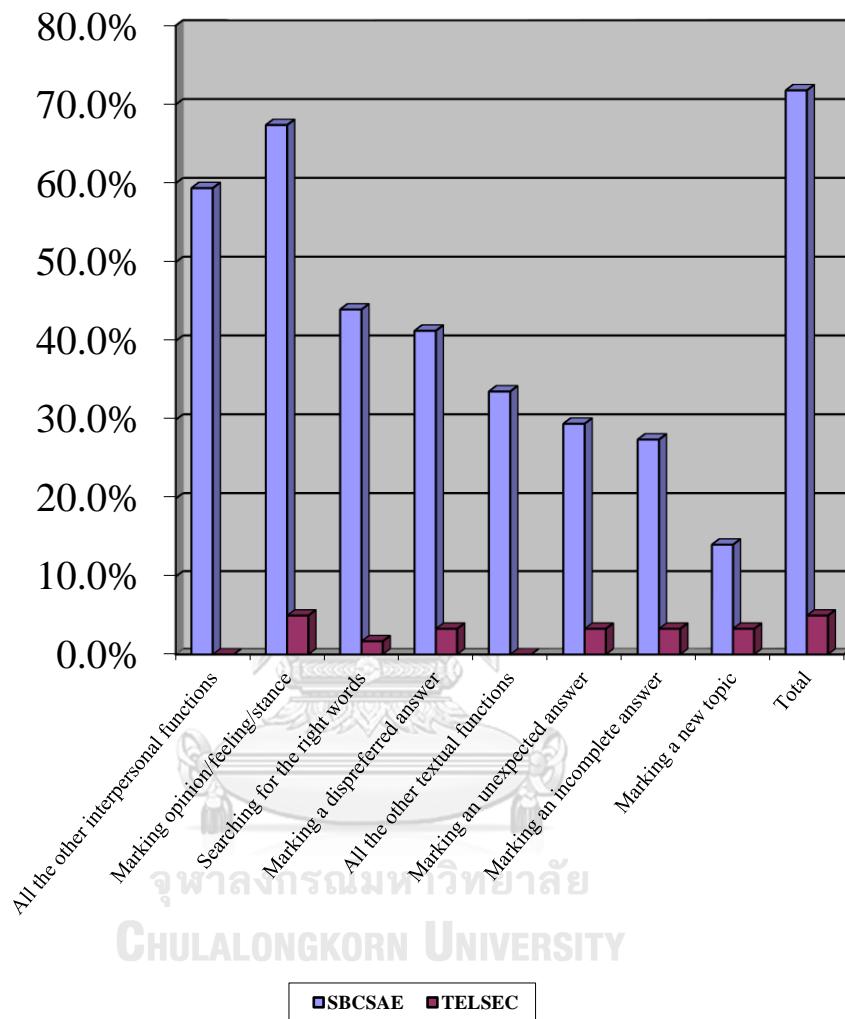
It is evidently to see the situation of the performance of *well* since seldom cases were found in TELSEC. From the table and two figures are shown above, Thai EFL learners who used the discourse marker *well* in this research only focused on a limited number of functions, while other functions, especially the interpersonal functions, were excluded from the learner corpus, including as a face-threat mitigator, marking a request, marking a question, etc., which were found in SBCSAE. In the meantime, a

very limited number of Thai EFL learners, with only 3 out of 60, performed *well*, resulting in a large difference in the proportion of the subjects in performing each function between Thai EFL learners and native English speakers.

Among the functions used by Thai EFL learners, it was found that the two functions, searching for the right words and marking opinion, feeling or stance, had the largest differences in terms of the frequencies compared to native English speakers, while only one function, marking a new topic, was used as the similar frequency as native English speakers.

The results pointed above directly show a lack of use of the spoken discourse marker *well* by Thai EFL learners in all the three facets that this research concerns. However, the result of this research is a little different from some previous research in which they focus on *well* as well (Aijmer, 2011; Müller, 2005). The previous research suggested that EFL learners used some functions of *well* more than American native English speakers, including searching for the right words and as a repair. However, these two functions were used in low frequencies by Thai EFL learners. Considering the fact that only 24 cases of using *well* and 3 C1-level participants were found in this research, it can be concluded that the discrepancy in using the discourse marker *well* is huge between native English speakers and Thai EFL learners.

A point worthy of mentioning concerning the discourse marker *well* is that it was believed that *well*, as a discourse marker in conversation, was known as its function to connect the relationship between the speaker and the hearer, resulting in its significant interpersonal functions that were worth studying than its textual functions illustrated by the previous research (Aijmer, 2011; Jucker, 2002; Sakita, 2013). Given the high frequency in the use of the interpersonal functions and high percentage of native English speakers who used different interpersonal functions in SBCSAE, it was in line with the previous research on the same point. Therefore, from the low

performance of *well* with few participants who used it, it can be seen that Thai EFL learners have a large deficiency in using the interpersonal functions of *well* to connect to the interlocutors in English conversation.

4.5.4 Different pragmatic functions of *you know* as a discourse marker

Table 48 illustrates the LL test results of each function and all the functions of *you know*, the raw frequency and its corresponding proportion of each function and all used functions of *you know*, the normalized frequency of each function and all used functions of *you know*, and the number and its corresponding proportion of the participants who used each function and all the functions of *you know* in both TELSEC and SBCSAE.

Table 48: Comparisons of each function and all functions of you know in SBCSAE and TELSEC

You know	SBCSAE			TELSEC			LL	SBCSAE		TELSEC	
Function	**RF**	**P (%)**	**NF**	**RF**	**P (%)**	**NF**		**NP**	**P (%)**	**NP**	**P (%)**
Marking an involvement (IF)	311	21.5	12	48	54.5	6	- 21.95	141	47.3	4	6.7
Marking opinion/feeling or stance (IF)	215	14.9	9	9	10.2	1	- 64.24	146	49.0	2	3.3
More explanation (TF)	173	12.0	7	4	4.5	1	- 64.73	166	55.7	2	3.3
Searching for the right words (TF)	168	11.6	7	1	1.1	0	- 79.30	129	43.3	1	1.7
All the other textual functions (TF)	121	8.4	5	0	0	0	- 55.16	64	21.5	0	0
All the other interpersonal functions (IF)	112	7.6	4	0	0	0	- 50.56	103	34.6	0	0
An implicature of common knowledge (IF)	97	6.7	4	23	26.1	3	- 1.14	77	25.8	4	6.7
Leading to an imaginary scene (TF)	79	5.5	3	2	2.3	0	- 28.77	79	26.5	1	1.7
Marking an emphasis/repetition (IF)	15	1.0	1	1	1.1	0	- 3.36	13	4.4	1	1.7
Total	1444	100	58	90	100	12	- 339.80	174	58.4	5	8.3
Total number of participants in the corpus								298	100	60	100

Figure 15 illustrates the proportion of each function of *you know* used by Thai EFL learners and native English speakers, while Figure 16 illustrates the proportion of the participants who used each function of *you know* used by Thai EFL learners and native English speakers.

Figure 15: Proportion of each function of the discourse marker you know used by Thai EFL learners and native English speakers
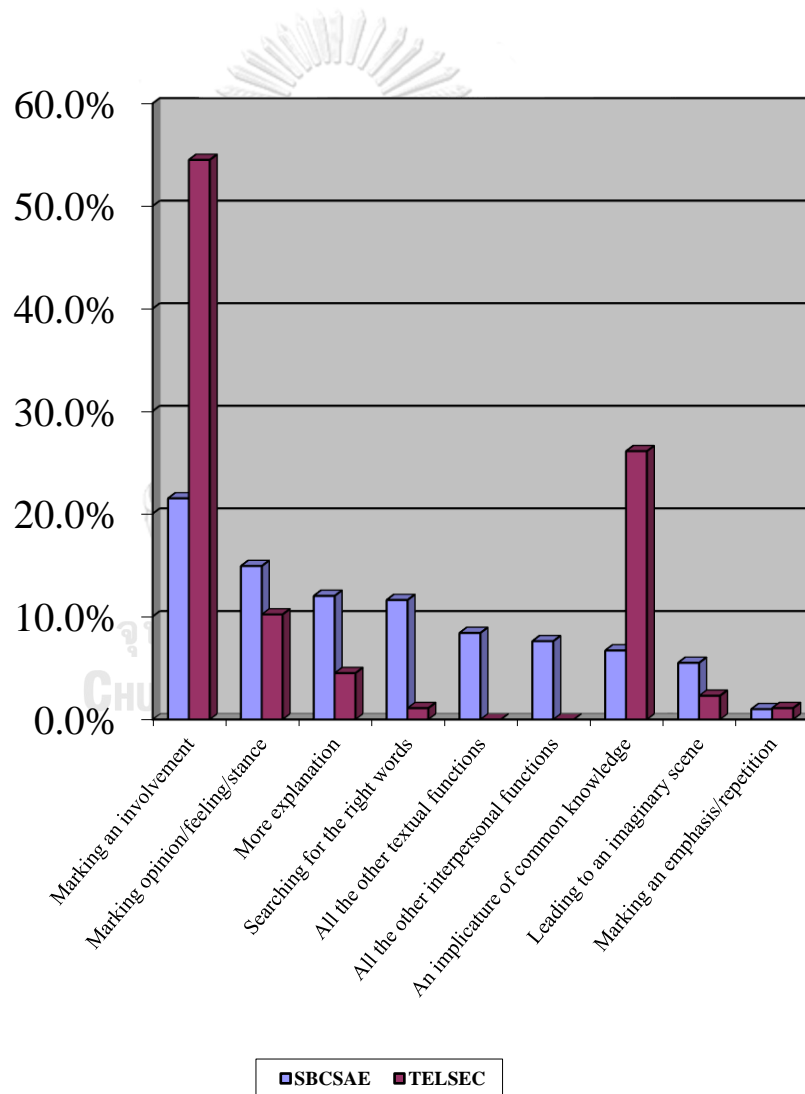
Figure 16: Proportion of participants who used the function of the discourse marker you know in TELSEC and SBCSAE

Similar to the situation of the discourse marker *well*, due to the low frequency of the performance of the discourse marker *you know*, it is observed that Thai EFL learners only focused on two functions of *you know*, while native English speakers performed it in a larger range of its functions. Meanwhile, more than half of native English speakers in SBCSAE used *you know* while only five Thai participants use it.

The same as the situation of *like* and *well*, it is interesting to observe that native English speakers performed the function, searching for the right words, remarkably higher with a much larger proportion of subjects, while only one case performed by one Thai participant was found in the learner corpus. Only one function used by Thai EFL learners, namely, an implicature of the shared or common knowledge, was performed at a similar frequency as native English speakers. Besides, even though Thai EFL learners, based on Figure 15, tended to focus on the function, to acquire an understanding or an involvement from the hearer, the frequency of using it was still significantly lower than native English speakers. It should be noted that according to the previous research (Hellermann & Vergun, 2007; Müller, 2005), two interpersonal functions, namely, an implicature of shared or common knowledge and to acquire an understanding or an involvement from the hearer, were regarded as the major functions of the discourse marker *you know*. It was illustrated by Hellermann and Vergun (2007) how the two functions had become the major functions of the discourse marker *you know* from the diachronic point of view. However, it can be seen, from the table and figures above, that native English speakers did not specifically use these two functions, while Thai EFL learners mainly focused on using them, indicating that Thai EFL learners tended to stick to using these two functions of *you know* in English conversation. Based on House's (2009) research on *you know*, it argued that for the speakers who used English as Lingua Franca, they did not often consider the interpersonal purpose of using *you know*, wherein *you know* was mostly used from the speaker's purpose, i.e., coherence in conversation, without considering the connection with the hearer. It is therefore suggested that Thai EFL learners may have a lack of the understanding of using the spoken discourse marker *you know* in the way that they did not use it for the coherence in utterance (textual functions), but they also did not use it for keeping the relationship with the hearer in conversation

244

(interpersonal functions). It is also noted that many functions of *you know* found in the previous research were not performed by Thai EFL learners but found in SBCSAE, i.e., as a repair, providing more explanation, leading a quote, and so forth. in its textual functions; as a hedge or a mitigator in its interpersonal functions.

4.5.5 Different pragmatic functions of *I think* as a discourse marker

Table 49 illustrates the LL test results of each function and all the functions of *I think*, the raw frequency and its corresponding proportion of each function and all used functions of *I think*, the normalized frequency of each function and all used functions of *I think*, and the number and its corresponding proportion of the participants who used each function and all the functions of *I think* in both TELSEC and SBCSAE.

Table 49: Comparisons of each function and all functions of I think in SBCSAE and TELSEC

I think	SBCSAE			TELSEC			LL	SBCSAE		TELSEC	
Function	RF	P (%)	NP	RF	P (%)	NP		NP	P (%)	NP	P (%)
An implicature of opinion/feeling/stance (IF)	340	68.3	14	408	91.7	54	+ 341.33	245	82.2	60	100
Searching for the following words (TF)	82	16.5	3	35	7.8	5	+ 2.80	80	26.8	9	15.0
A hedge/mitigator (IF)	76	15.3	3	2	0.4	0	- 27.34	61	20.5	2	3.3
Total	498	100	20	445	100	59	+ 259.30	251	84.2	60	100
Total number of participants in the corpus								298	100	60	100

246

Figure 17 illustrates the proportion of each function of *I think* used by Thai EFL learners and native English speakers, while Figure 18 illustrates the proportion of the participants who used each function of *I think* used by Thai EFL learners and native English speakers.

Figure 17: Proportion of each function of the discourse marker I think used by Thai EFL learners and native English speakers

Figure 18: Proportion of participants who used the function of the discourse marker I think in TELSEC and SBCSAE

It is overtly to see, from the table and figures above, that the function, namely an implicature of opinion, feeling or stance, had the largest proportions of participants who used this function in both corpora. However, the over-reliance on using this function was still found in the learner corpus TELSEC compared to the use of it by native English speakers in SBCSAE.

Unlike native English speakers in SBCSAE, Thai EFL learners performed the other two functions in much lower distributions with smaller proportions of participants, respectively. It is noted that for the function, searching for the following words, Thai EFL learners used it with a lower proportion of participants than native English speakers. Moreover, the function, as a mitigator, was underused by Thai EFL learners based on all the three facets examined in this research.

The result of using *I think* by Thai EFL learners indicates that Thai EFL learners mainly focused on one particular function of *I think*, while native English speakers used the other two functions of *I think* in a larger amount of participants in SBCSAE, with a much higher frequency in one of the functions, as a mitigator. It suggests that Thai EFL learners stuck to using one function of the spoken discourse marker *I think*, which leads to inappropriate use of *I think* in English conversation.

4.5.6 Different pragmatic functions of *I mean* as a discourse marker

Table 50 illustrates the LL test results of each function and all the functions of *I mean*, the raw frequency and its corresponding proportion of each function and all used functions of *I mean*, the normalized frequency of each function and all used functions of *I mean*, and the number and its corresponding proportion of the participants who used each function and all the functions of *I mean* in both TELSEC and SBCSAE.
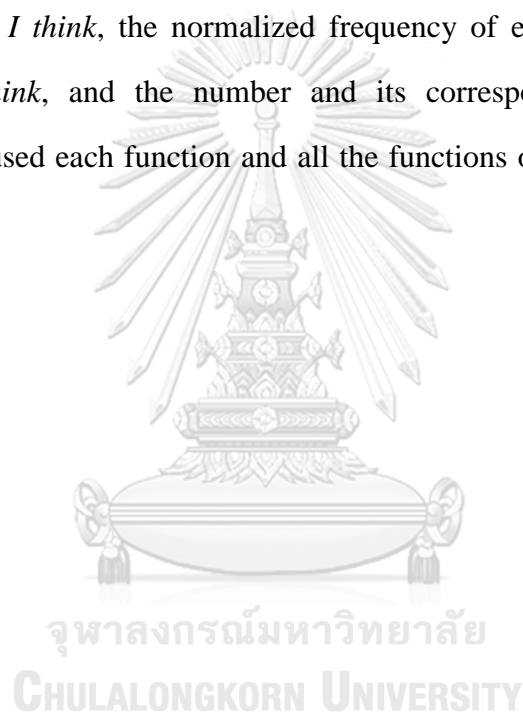
Table 50: Comparisons of each function and all functions of I mean in SBCSAE and TELSEC

I mean	SBCSAE			TELSEC			LL	SBCSAE		TELSEC	
Function	RF	P (%)	NP	RF	P (%)	NP		NP	P (%)	NP	P (%)
An implicature of the cause subjectively (IF)	124	22.6	5	15	12.1	2	- 14.15	115	38.6	4	6.7
As a mitigator (IF)	92	16.8	4	0	0	0	- 40.40	88	29.5	0	0
Marking opinion/feeling/stance (IF)	84	15.3	3	28	22.6	4	+ 0.20	76	25.5	7	11.7
More explanation (TF)	80	14.6	3	44	35.5	6	+ 9.53	75	25.2	7	11.7
Searching for the right word (TF)	48	8.8	2	1	0.8	0	- 18.48	45	15.1	1	1.7
Emphasis/repetition (IF)	40	7.3	2	3	2.4	0	- 8.11	39	13.1	1	1.7
As a repair (TF)	20	3.6	1	28	22.6	4	+ 27.20	17	5.7	4	6.7
Marking a topic shift (TF)	20	3.6	1	5	4.0	1	- 0.15	11	3.7	3	6.7
Total	546	100	22	124	100	16	- 8.90	156	52.3	15	25.0
Total number of participants in the corpus								298	100	60	100

Figure 19 illustrates the proportion of each function of *I mean* used by Thai EFL learners and native English speakers, while Figure 20 illustrates the proportion of the participants who used each function of *I mean* used by Thai EFL learners and native English speakers.

Figure 19: Proportion of each function of the discourse marker I mean used by Thai EFL learners and native English speakers
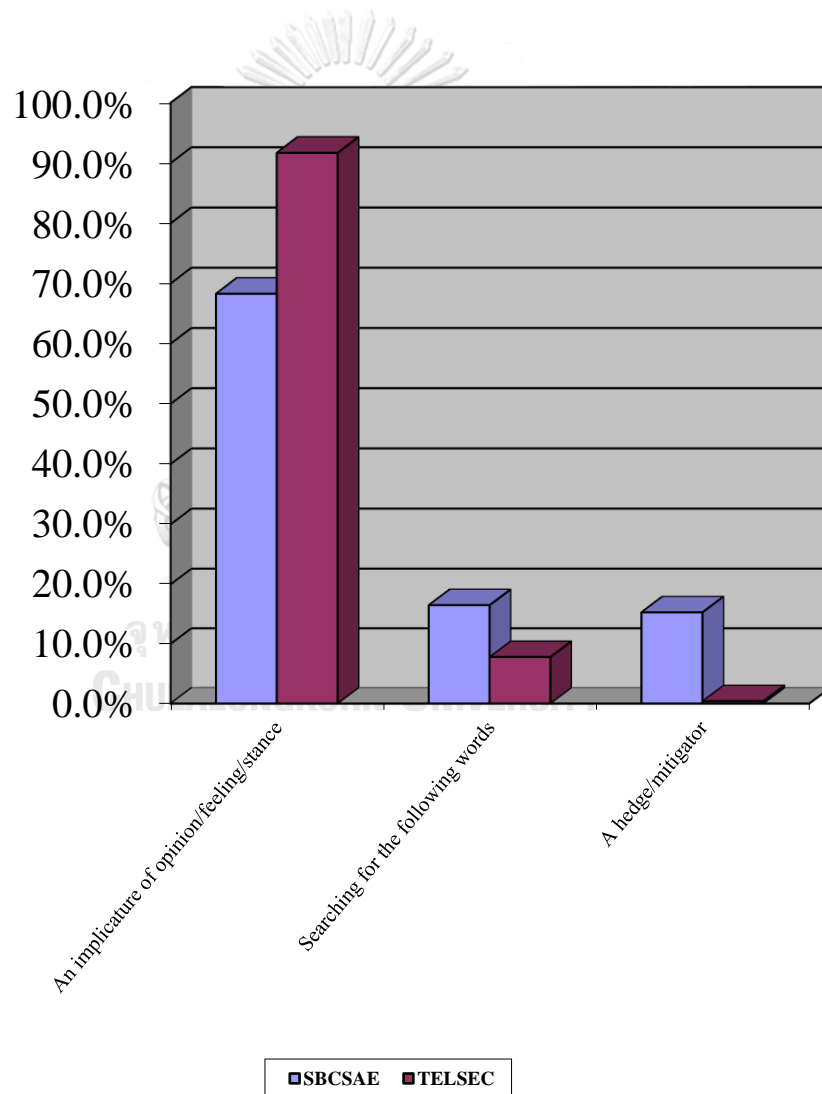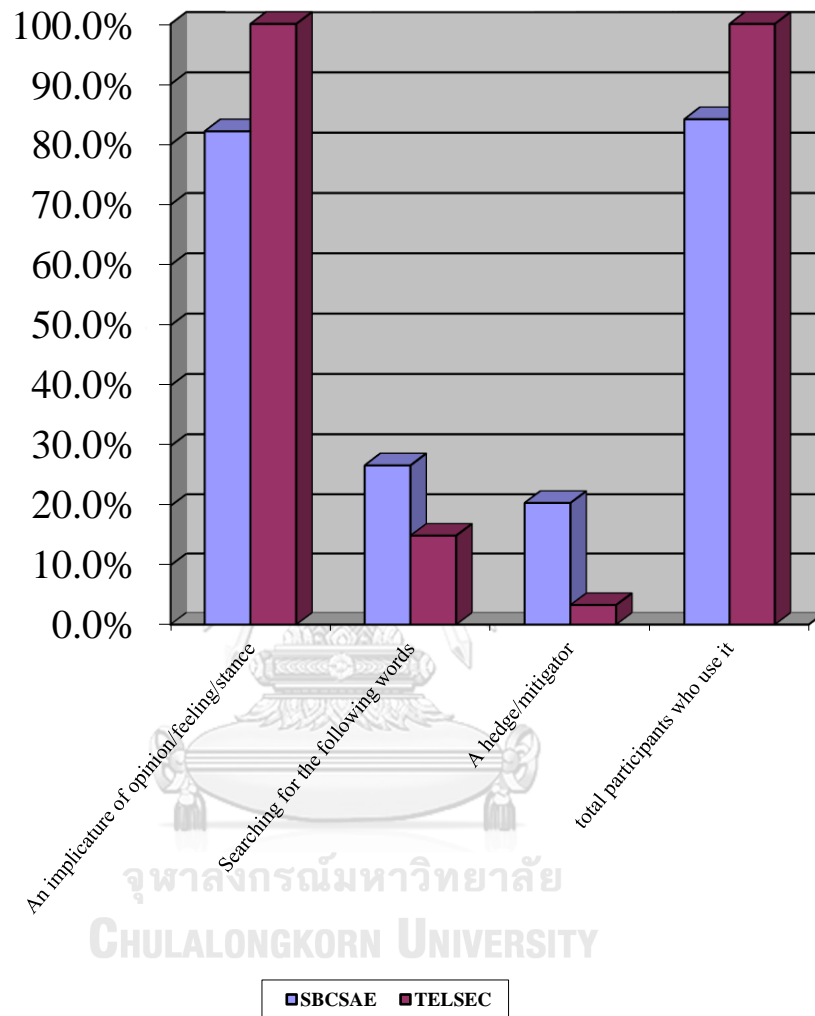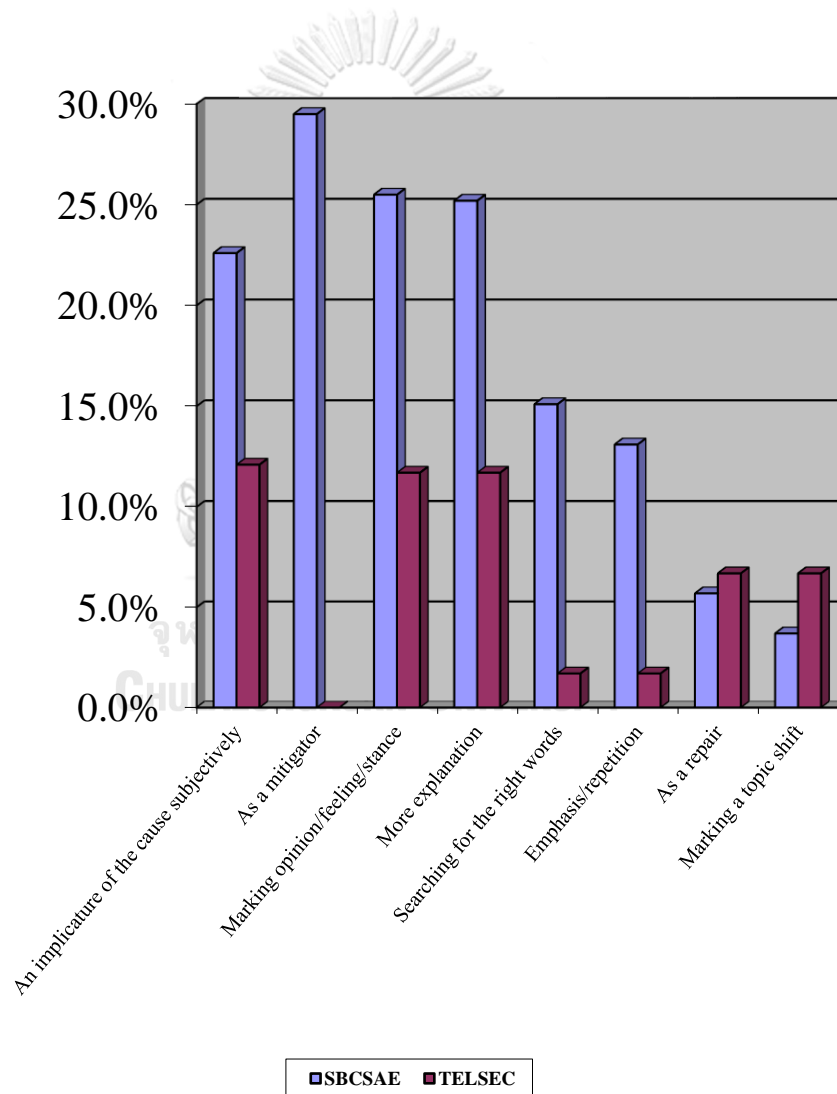
Figure 20: Proportions of participants who used the function of the discourse marker
I mean in TELSEC and SBCSAE
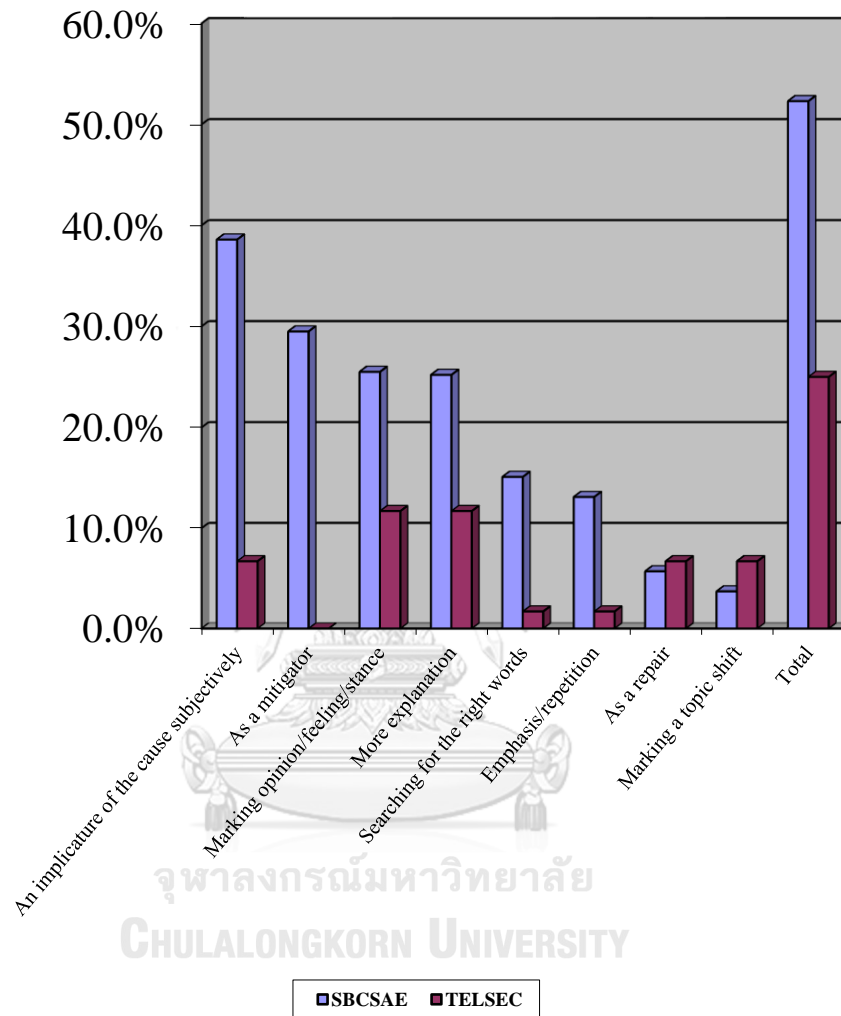
Based on the table and figures shown above, the use of *I mean* between two corpora was different, although there was no significant difference according to the answer to question 1. Native English speakers used the function, an implicature of the cause subjectively, in the highest frequency, while Thai EFL learners used the function, more explanation, in the highest frequency. No performance of the function, as a mitigator, was found in TELSEC, whereas native English speakers used this

function in a high frequency with a large number of participants. The function, searching for the following words, was also underused by Thai EFL learners in the discourse marker *I mean*, which was the same condition in the discourse marker *like*, *well* and *you know*. Another function, as a repair, as was argued as one of the derived functions from its semantic meaning (Brinton, 2010), was used much more often by Thai EFL learners than native English speakers. Moreover, the proportion of Thai participants who use *I mean* was half of native English speakers, with large differences in six functions out of the total eight functions.

As is illustrated by the data, it shows that Thai EFL learners mainly focused on two textual functions and one interpersonal function of *I mean*, namely, more explanation, as a repair and marking opinion, feeling or stance, whereas native English speakers mainly used it in three interpersonal functions, namely, an implicature of the cause subjectively, as a mitigator and marking opinion, feeling or stance. Based on previous research (Brinton, 2008, 2010), the two textual functions that Thai EFL learners focused on were derived from the original semantic meaning of *I mean*. It may indicate that it is easier for Thai EFL learners to perform the functions that are close to the original semantic meaning of a spoken discourse marker. The same as the use of previous spoken discourse markers, the proportion of Thai participants who used *I mean* suggests that only a few Thai EFL learners may have an understanding of using the spoken discourse markers in English conversation.

In the next part, this research will illustrate different uses of spoken discourse markers in a whole picture between Thai EFL learners and native English speakers.

4.5.7 A conclusion of different pragmatic functions used by Thai EFL learners compared to native English speakers

After the analysis of different uses of pragmatic functions of each discourse marker between the two corpora, this part presents different uses of the pragmatic

functions by the observation of all the six spoken discourse markers between Thai EFL learners and native English speakers as a whole picture.

To illustrate the differences as a whole picture, this research mainly presents six facets: differences in the use of types of the textual functions and the interpersonal functions, differences in using all the textual functions, differences in using all the interpersonal functions, differences in the number of participants who used all the textual functions, differences in the number of participants who used all the interpersonal functions, and two detailed differences in using the functions that were shared by multiple spoken discourse markers in this research. The analysis is presented as follows.

Based on the data in both corpora, 39 pragmatic functions were found in the learner corpus TELSEC, while 56 pragmatic functions were found in the native English speakers corpus SBCSAE. It can be seen that Thai EFL learners used 70% of pragmatic functions compared to native English speakers, as illustrated in Table 51 below.

Table 51: Number of textual functions and interpersonal functions of the discourse markers used in each corpus

Discourse marker	Number of textual functions used in each corpus		Number of interpersonal functions used in each corpus	
	TELSEC	SBCSAE	TELSEC	SBCSAE
like	6	6	0	0
so	6	8	4	4
well	2	5	4	9
you know	2	6	5	7
I think	1	1	2	2
I mean	4	4	3	4

The total	21	30	18	26

It should reiterate, as has been pointed out in the previous part, that 9 out of 39 pragmatic functions used by Thai EFL learners were found to have only one to three instances in TELSEC. Considering the generalization in the use of the spoken discourse markers, 30 out of 39 pragmatic functions were regarded as generally used pragmatic functions by Thai EFL learners.

Table 52 illustrates the LL test results of each type of function and both types of functions of all the six spoken discourse markers, the raw frequency and its corresponding proportion of each type of function and both types of functions of all the six spoken discourse markers, the normalized frequency of each type of function and both types of functions of all the six spoken discourse markers, and the number and its corresponding proportion of the participants who used each type of function and both types of functions of all the six spoken discourse markers in both TELSEC and SBCSAE.`

Table 52: Comparison of both types of functions between SBCSAE and TELSEC

Total six spoken discourse markers	SBCSAE			TELSEC			LL	SBCSAE		TELSEC	
Function	RF	P (%)	NP	RF	P (%)	NP		NP	P (%)	NP	P (%)
All the textual functions	3391	49.9	136	689	53.4	92	- 97.65	257	86.2	32	53.3
All the interpersonal functions	3400	50.1	137	602	46.6	80	- 164.22	265	88.9	60/14	100/23.3
Total	6791	100	273	1291	100	172	- 256.75	271	90.9	60/34	100/56.7
Total number of participants in the corpus								298	100	60	100

It is worth mentioning here a point concerning the number of participants who used interpersonal functions in TELSEC, as also pointed out in the previous part. 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 not appropriate 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.

Overall, as is illustrated by Table 50, it is observed that Thai EFL learners performed 70% of both textual functions and interpersonal functions of spoken discourse marker compared to the use of native English speakers. This result indicates that Thai EFL learners have not fully understood the use of all the functions of spoken discourse markers in conversation.

Based on Table 51, it is interesting to find that native English speakers used both textual functions and interpersonal functions almost 50%, respectively, while Thai EFL learners used textual functions slightly more than interpersonal functions. However, as the LL results show, Thai EFL learners used both types of functions significantly different from native English speakers, indicating a deficiency in performing spoken discourse markers in both types of functions by Thai EFL learners. Moreover, as the illustration below in table 51, if the interpersonal function of *I think*, namely, marking opinion, feeling and stance, is taken out of consideration, only 56.7% of Thai EFL learners used spoken discourse markers, with 53.3% using textual functions and 23.3% using interpersonal functions, while 90.9% native English speakers used spoken discourse markers, with 86.2% used textual functions and 88.9% used interpersonal functions. As the data shows, Thai EFL learners used interpersonal functions fewer than textual functions, and fewer participants used interpersonal functions than textual functions. It is therefore conceivable that the discrepancy in the performance of the interpersonal functions of the spoken discourse markers by Thai EFL learners is larger than the performance of the textual functions, i.e., a relatively

low performance of the interpersonal functions of *so*, *well*, *you know* and few participants who produced each interpersonal function in different spoken discourse markers compared to native English speakers.

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. Similar are the findings in Arya (2020) and Nookam (2010) wherein a general deficiency in using spoken discourse markers in English conversation was found. Moreover, comparisons of functions of spoken discourse markers in this research have reinforced the notion, as similarly pointed out in Arya (2020) and Nookam (2010), that Thai EFL learners lacked 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. However, 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 reasons to consider the influence of L1 Thai regarding lack of use of English spoken discourse markers by Thai EFL learners. With previous research in mind, it was found that native Thai speakers used Thai spoken discourse markers for politeness and intimacy (Kittopakrankit, 2018) rather than for interpersonal purposes, with cultural and social impact a significant factor here. Moreover, as previously suggested (Chotiros, 1999), discourse markers in both languages may have no one-to-one correspondences, making it difficult for Thai EFL learners to acquire English spoken discourse markers.

After the results and the discussion on the use of pragmatic functions of spoken discourse markers by Thai EFL learners in a whole picture, this research further discusses two details that were found unique based on the data of question 4.

First, Thai EFL learners tended to use the functions that were derived from the original semantics of the discourse markers, for instance, the high frequency in the use of the function of the discourse marker *like*: to exemplify; the function of the discourse marker *so*: marking a result or consequence; the function of the discourse

258

marker *I think*: an implicature of personal opinion, feeling or stance; the function of the discourse marker *I mean*: marking more explanation or exemplification, etc., or that were most commonly used functions found in previous research, for instance, the high frequency of the use of the function of the discourse marker *well*: marking opinion, feeling or stance; the function of the discourse marker *you know*: to acquire an understanding or an involvement from the hearer, etc. The frequencies of these functions listed above used by Thai EFL learners were at least twice higher than the frequencies of them used by native English speakers. It was observed that native English speakers also performed these functions in high frequencies, but not so high as Thai EFL learners. It reflects that, on the one hand, it is easier for EFL learners to acquire the functions of the discourse markers that are similar to or derived from their original semantic meanings (Brinton, 2008) so that they stick to using them in conversation; on the other hand, it also indicates, as found in previous research (Asik & Cephe, 2013; Fung & Carter, 2007), that some functions of the discourse markers are considered more difficult for EFL learners to acquire.

Upon a careful examination of each spoken discourse marker in this research, it is interesting to find that four functions, shared by different spoken discourse markers, were commonly performed in low frequencies by Thai EFL learners, which were found in both textual functions and interpersonal functions.

First of all, it is interesting to find that the function, searching for the following words, was performed much less frequently with a small number of participants who used it in different spoken discourse markers such as *so*, *well*, *you know* and *I mean*. Table 53 illustrates the raw frequencies and proportions of using this function in five different spoken discourse markers, the normalized frequencies and the LL test results of this function in each spoken discourse marker, and the number and proportions of the participants who used it in five spoken discourse markers in both corpora.

Table 53: Comparisons of using the function, searching for the right words, in five spoken discourse markers between the two corpora

Searching for the right words					
	like	*well*	*You know*	*I think*	*I mean*

RF in SBCSAE	380	214	168	82	48
P (%)	25.9	15.3	11.6	16.5	8.8
NF	15	9	7	3	2
RF in TELSEC	15	2	1	35	1
P (%)	3.5	8.3	1.1	7.8	0.8
NF	2	0	0	5	0
LL	- 116.77	- 95.36	- 79.30	+ 2.80	- 18.42
NP in SBCSAE	205	133	129	100	53
P (%)	68.8	44.6	43.3	33.6	17.8
NP in TELSEC	4	1	1	2	1
P (%)	6.7	1.7	1.7	3.3	1.7

It was found that this function assisted the speaker to buy more time for the purpose of considering the next utterance at the moment in conversation (Fung & Carter). It was discovered that Thai EFL learners used fillers, such as *emm*, *uhhm*, etc., to buy more time thinking of the next utterance they intended to express, as illustrated by the two excerpts found in TELSEC.

> INT: And (.) you don't like it ?
> Mr. B1-14: (.) &-emm (..) &-uhh (..) I don't (.) +/ .
> Mr. B1-14: +/ not hate it (.) but sometime feel not good .
> (Mr. B1-14)

Excerpt 93

> Ms. C1-12: I see it in a (.) different way .
> INT: How different ?
> INT: What do you mean ?
> Ms. C1-12: Oh (..) .
> Ms. C1-12: <laugh> .
> Ms. C1-12: &-uhh (.) I [/] &-emm (..) .
> Ms. C1-12: &-uhh (.) ok I once was in a park .
> Ms. C1-12: Just (.) waiting for someone .

(Ms. C1-12)

Excerpt 94

In both excerpts above, the participants insert some fillers, such as *emm*, *uhh*, etc., to buy more time to think of the next utterance. Due to the redundancy of different fillers, the conversations, as shown in both excerpts, appear to be inconsistent and incoherent, which may result in impatience or misunderstanding from the hearer's point of view.

As commonly used by native English speakers, instead of the repetition of the fillers, different spoken discourse markers in various contextual environments signal the purpose of searching for the following words for the speaker. Therefore, instead of using fillers, if the participants in both excerpts above insert the spoken discourse marker *well* or *you know*, the redundancy of using fillers can be avoided, and the problem of inconsistency and incoherence of the conversations can be appropriately solved.

As the previous research (Asik & Cephe, 2013) suggested, non-native English speakers had the fossilization of using the fillers in case of searching for the words they needed to speak in the next utterance. They did not or seldom used the spoken discourse markers to do so because they may lack the awareness of the existence of this function that various spoken discourse markers can offer in conversation.

Another textual function that Thai EFL learners rarely used by different spoken discourse markers was as a repair, shared by *so*, *well*, *you know*, etc. Based on the previous research (Aijmer, 2016; Brinton, 2008; Diskin, 2017), the speaker realizes an error in the preceding utterance so that he or she inserts a spoken discourse marker to signal the hearer about it and does a self-correction after the discourse marker. Table 54 illustrates the raw frequencies and proportions of using this function in five different spoken discourse markers, the normalized frequencies and the LL test result of this function in each spoken discourse marker, and the number and proportions of the participants who used it in five spoken discourse markers in both corpora.

Table 54 Comparisons of using the function, as a repair, in five spoken discourse markers between the two corpora

As a repair					
	like	*so*	*well*	*you know*	*I mean*
RF in SBCSAE	8	18	54	51	20
P (%)	0.5	1.0	4.0	3.5	3.6
NF	0	1	2	2	1
RF in TELSEC	0	0	0	0	28
P (%)	0	0	0	0	22.6
NF	0	0	0	0	4
LL	- 18.46	- 9.50	- 28.49	- 26.90	+ 27.20
NP in SBCSAE	10	7	44	37	17
P (%)	3.4	2.3	14.8	12.4	5.7
NP in TELSEC	0	0	0	0	4
P (%)	0	0	0	0	6.7

As the data presents, native English speakers performed this function in various spoken discourse markers investigated in this research while Thai EFL learners excluded this function except for the discourse marker *I mean*. Like the function, searching for the following words, this function was considered a textual function for the speaker to correct the error in the preceding utterance (Aijmer, 2011).

It is therefore assumed that Thai EFL learners did not have the awareness of this function shared by different spoken discourse markers so that they were not competent in performing it in oral communication. The excerpt below may present a situation where the participant can insert a spoken discourse marker for the purpose of this function.

> Ms. B1-24: (.) yes (.) maybe they can .
>
> Ms. B1-24: But (.) I'm not [/] not sure of it .
>
> Ms. B1-24: (.) because they have no money .
>
> Ms. B1-24: (.) &-uhh (.) little money (.) not many (.) .
>
> (Ms. B1-24)

Excerpt 95

The participant in excerpt 95 has a self-correction from "no money" to "little money" with an additional explanation "not many". Instead of using the filler *uhh*, the spoken discourse marker *well* or *I mean* can be used. It is thus a strategy to use spoken discourse markers in this function to signal a self-correction in conversation.

In the interpersonal functions, it is noticed that as a mitigator, one of the important interpersonal functions that was found in SBCSAE and in the previous research (Aijmer, 2011, 2016; Bolden, 2015; Brinton, 1996, 2008; Diskin, 2017), was almost excluded by Thai EFL learners. In SBCSAE, it is shared by the discourse marker *well*, *you know, I think* and *I mean*. Diskin (2017) also discovered this function in the discourse marker *like* used by native English speakers. Only two cases of using this function were found in the discourse marker *I think* in TELSEC, reflecting the fact that Thai EFL learners did not perform this function in English conversation. Table 55 illustrates the raw frequencies and proportions of using this function in five different spoken discourse markers, the normalized frequencies and the LL test result of this function in each spoken discourse marker, and the number and proportions of the participants who used it in four spoken discourse markers in both corpora.

Table 55: Comparisons of using the function, as a mitigator, in four spoken discourse markers between the two corpora

As a mitigator				
	well	*you know*	*I think*	*I mean*
RF in SBCSAE	274	100	76	92
P (%)	19.7	6.9	15.3	16.8
NF	11	4	3	4
RF in TELSEC	0	0	2	0
P (%)	0	0	0.4	0
NF	0	0	0	0
LL	- 194.07	- 59.08	- 27.34	- 40.40
NP in SBCSAE	152	89	61	88
P (%)	51.0	29.9	20.5	29.5

NP in TELSEC	0	0	2	0
P (%)	0	0	3.3	0

Based on the previous research (Aijmer, 2004, 2011; Asik & Cephe, 2013; Buysse, 2012; Dehe & Wichmann, 2010; Diskin, 2017; Fitzmaurice, 2004; Fung & Carter; 2007; Müller, 2005; Sakita, 2013; Tree, 2010; Vanda & Peter, 2011), this function was discovered at the moment when a potential argument appeared or when the speaker tried to ease the tension with various reasons, such as a negative response, a refuse to the preceding utterance, a disagreement, etc. With certain tones or longer pauses around, the spoken discourse marker in this type of context functions as a mitigator or a hedge to specifically signal the hearer in which the speaker intends to relieve the potential argument or the tension at the moment in conversation. It is thereby associated with the good intention of the speaker as such. From this point of view, it is overtly to see that this function is highly relevant to the relationship between the speaker and the hearer from the interpersonal perspective, in which the signal that the speaker sent for mitigating the conversational atmosphere by using the spoken discourse marker in context should be achieved by the hearer who understands the signal in the same sense so that he or she can feel the intention of the speaker. In SBCSAE, it was found that this function that was shared by four spoken discourse markers listed in the table above was used relatively in high frequencies and was performed by more than 20% of native English speakers in each spoken discourse marker. However, almost no case of using this function was found by Thai EFL learners. Even though two cases were found in the discourse marker *I think*, the LL test result showed a significant difference.

There may have chances for Thai EFL learners to perform this function in conversation by using some spoken discourse markers, as illustrated by the excerpt below.

> INT: (.) oh (.) so [DM] you don't think it's ok ?
> Ms. C1-18: no (.) I don't think it's ok .
> Ms. C1-18: I totally disagree with that .
> INT: oh ok (.) but they need to make money .

Ms. C1-18: <laugh> .

Ms. C1-18: but I'm sure there's other way .

Ms. C1-18: my thinking of course .

(Ms. C1-18)

Excerpt 96

There are two tensions appearing in the excerpt above. The first tension appears when the participant disagrees with the hearer by stating "no, I don't think it's ok". The second tension appears when the participant argues against a previous statement "but they need to make money". It is overtly to see that the participant directly presents her disagreement or argument against the statements by the hearer. As is mentioned earlier, it may not cause the understanding semantically, but may cause a problem pragmatically in that the hearer is disagreed twice consecutively and directly where the hearer may be exposed to a face-threat or an uncomfortable emotion from the speaker's disagreement and argument. It is thus appropriate to insert some discourse marker, such as *well*, *you know* or *I mean* to ease the tension or the face-threat problem.

As this function highly requires the knowledge and the awareness of the speaker to understand how to perform it when the potential argument or the tension appears, the result suggests that it should be considered a difficult pragmatic function for Thai EFL learners to achieve and perform in English conversation. As being one of the highly used interpersonal functions by native English speakers, it may indirectly reflect the deficiency in the use of interpersonal functions of the spoken discourse markers by Thai EFL learners.

Another interesting finding in the interpersonal functions is marking opinion, feeling or stance shared by the discourse marker *so*, *well*, *you know*, *I think* and *I mean*. Table 56 illustrates the raw frequencies and proportions of using this function in five different spoken discourse markers, the LL test result of this function in each spoken discourse marker, and the number and proportions of the participants who use it in four spoken discourse markers in both corpora.

Table 56: Comparisons of using the function, marking personal opinions, feelings or stance, in five spoken discourse markers between the two corpora

Marking personal opinions, feelings or stance					
	so	*well*	*You know*	*I think*	*I mean*
RF in SBCSAE	274	252	215	340	84
P (%)	15.7	18.1	14.9	68.3	15.3
NF	11	10	9	14	3
RF in TELSEC	39	7	9	408	28
P (%)	20.7	29.2	10.2	91.7	22.6
NF	5	1	1	54	4
LL	- 23.18	- 89.04	- 64.24	+ 341.33	+ 0.20
NP in SBCSAE	146	201	146	245	76
P (%)	49.0	67.4	49.0	82.2	25.5
NP in TELSEC	7	3	2	60	7
P (%)	11.7	5.0	3.3	100	11.7

If only examining the data in TELSEC, it reveals that Thai EFL learners used this function in high frequencies, with higher distributions of this function in each discourse marker. However, compared to the data in SBCSAE, it was found that this function was generally underused in the discourse marker *so*, *well* and *you know* while it was overused in the discourse marker *I think*. The distributions of the Thai EFL learners who performed this function in the discourse marker *so*, *well you know* and *I mean* were also much lower than native English speakers in SBCSAE. As illustrated by the table above, it reflects that Thai EFL learners should be guided to use this function with the discourse marker *so*, *well* and *you know*. Meanwhile, they should also avoid the over-reliance on using the discourse marker *I think.* This finding indicates that Thai EFL learners need further attention to this function. This is because, on the one hand, its performance in frequencies in different spoken discourse markers is not as similar as the use of native English speakers; on the other hand, much fewer subjects perform this function compared to native English speakers. It is therefore considered that Thai EFL learners have not mastered this function completely.

It should be noted concerning the comparisons of the differences in the pragmatic functions used by Thai EFL learners and native English speakers in this research that this comparison was based on the differences between the learner corpus TELSEC built in this research and the native English speaker corpus SBCSAE that shared the same genre of the spoken English data. Some functions that were discovered in the previous research were not found in both TELSEC and SBCSAE due to the use of different corpora, methodology or theoretical frameworks, i.e., the function: as a mitigator in the discourse marker *like* (Diskin, 2017); the functions: marking an agreement, marking a reinforcement, etc., in the discourse marker *well* (Aijmer, 2011; Fung & Carter, 2007). Since this research focused on the comparison between the learner corpus TELSEC and the native English speaker corpus SBCSAE, the functions that were found in the previous research but not found in both corpora are not further discussed here.

Besides all the manifestations for showing the lack of pragmatic competence in communication and incomplete English acquisition in the discussion above, this research provides three more points worth discussing: low performance of *well* and *you know*; over-reliance of *I think* and *like*; unbalanced performance of each function in a spoken discourse marker.

It was found that some spoken discourse markers mainly served their functions for interpersonal purposes in communication, such as the discourse marker *well* and *you know*, as the interpersonal functions were demonstrated in previous research (Aijmer, 2011; Blakemore, 2002; Hellermann & Vergun, 2007; Müller, 2005; Polat, 2011; Sakita, 2013; Trillo, 2002), leading to a result that both discourse markers are versatile in terms of their interpersonal functions to closely forge a bond between interlocutors. Therefore, it can be interpreted that *well* and *you know* are considered as interpersonal-centered spoken discourse markers in communication. Thanks to their multi-faceted interpersonal functions that only exist and can be studied in communication, it is more difficult for non-native English speakers to acquire them and use them appropriately (Polat, 2011). It is thus conceivable to notice that Thai EFL learners performed these two spoken discourse markers not only in a much lower frequency than others but also in a quite large significant difference compared to

native English speakers. This phenomenon, the rare performance of the interpersonal-centered spoken discourse markers, evidently reveals the condition where Thai EFL learners lack pragmatic competence. Furthermore, it indicates that seldom chances may be exposed to Thai EFL learners for the acquisition of these two spoken discourse markers in communication, resulting in a deficiency of being a competent English speaker from pragmatic perspective.

On the other hand, previous research also proposed the concept of pragmatic fossilization (Trillo, 2002), stating that like the traditional concept of fossilization in L2 or foreign language learning, pragmatic fossilization refers to the inappropriateness or unacceptability at the pragmatic level of communication by non-native speakers. Moreover, it was illustrated that the inappropriateness of the discourse markers led to pragmatic fossilization (Asik & Cephe, 2013; Trillo, 2002). Besides the incompetence or the low competence of using discourse markers in communication mentioned above, the pragmatic fossilization also includes the overuse of the discourse markers, no progress in using discourse markers with the increase of the English level, etc. It is therefore associated with the unusual performance of some discourse markers by Thai EFL learners compared to native English speakers, as pointed out above, with the over-reliance on the discourse marker *I think* by both Thai EFL learners and *like* by Thai C1-level EFL learners. This phenomenon of an over-reliance on some discourse markers, as also found in previous research (Asik & Cephe, 2013; Fung & Carter, 2007), indicates a fossilized focus on certain discourse markers.

A closer investigation of the LL test result of each spoken discourse marker between Thai EFL learners and native English speakers leads to the third manifestation, in which the LL test results of the two discourse markers that did not have a significant difference, namely, *like* and *I mean*, cannot be interpreted that Thai EFL learners performed both spoken discourse markers the same as native English speakers. Given the data shown in the previous part, it can be seen that the frequencies of three functions of the discourse marker *like* performed by Thai EFL learners were not as similar as the use of native English speakers. Meanwhile, the frequencies of three functions of *I mean* performed by Thai EFL learners were not as

similar as the use of native English speakers. As has been stated above, supported by previous research (Diskin, 2017; Polat, 2011; Sorace, 2004), the use of some discourse markers, concerned with their functions, positions, social or cultural influence, requires learners' high ability of language acquisition. This unbalanced performance of each function of the spoken discourse markers reflects the instability in the acquisition of the spoken discourse markers.

Overall, the discussions above reveal a discrepancy in the performance of the spoken discourse markers between Thai EFL learners and native English speakers in conversation. Furthermore, this discrepancy, to some extent, as demonstrated by the frequency and the LL test result, should not be ignored, i.e. the general underuse of functions in the discourse marker *so*, *well*, *you know*, etc.; the low distributions of Thai EFL learners who performed various discourse markers; the rare performance of *well* and *you know*; no significant improvement of the performance of *you know* from B1-level to C1-level; the over-reliance of *I think*, etc. All these findings may lead to one significant issue in the relevance of using the spoken discourse markers in English conversation by Thai EFL learners: the deficiency in pragmatic competence in oral communication.

As is known, pragmatic competence, as significantly important as grammatical competence (Canale, 1983; Canale & Swain, 1980), is an indispensable component in oral communication. It is thus no doubt that a competent speaker (Fung & Carter, 2007) should not only convey the right semantic meanings but also master the pragmatic competence in different dialogic activities, including illocutionary competence, social linguistic competence, etc. (Canale, 1988). Upon the examination of discourse markers by previous research (Aijmer, 2011; Fung & Carter, 2007; House, 2013), it is believed that spoken discourse markers, as the multi-functionality of this whole category affects oral communication from the pragmatic perspective in various aspects, i.e., interpersonal relationship of interlocutors, situational understanding between the speaker and the hearer, etc., are regarded as a focal component of expressing the pragmatic meaning by native English speakers and an important acquisition for EFL learners to become a competent English speaker. Fung and Carter (2007) believed that an improvement in performing the spoken discourse

markers in communication can strengthen EFL learners' pragmatic competence, indicating the vital connection between the spoken discourse markers and pragmatic competence. It was also defined by House (2009) as pragmatic fluency that a display of appropriate pragmatic competence by using the discourse markers to master smooth continuity in communication (House, 2013). Hence, a proper performance of the spoken discourse marker, as Fung and Carter (2007) illustrated, reflects a speaker's ability of pragmatic competence by using the language from cultural, social and situational perspectives.

Tracing back, then, to the 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 discrepancy in the performance of spoken discourse markers compared to native English speakers. Interpretation further reveals, following 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 in previous research there does not inhere 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. However, as the interpretations above state and 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).

It should be noted again that all the findings and discussions above are based on the comparison between the learner corpus TELSEC built by this research and the native English speaker corpus SBCSAE that is comparable to this research. It is therefore predictable that there may be different results in the comparison of different native English speaker corpora in other spoken genres or other English dialects.

Further studies are encouraged to continue pursuing the comparisons by using a different native English speaker corpus.

Chapter 5
Conclusion

This research examined the use of the English spoken discourse markers by Thai EFL learners in daily conversation. This research focused on six spoken discourse markers that have been commonly investigated in the use of native English speakers in previous research but have not been examined comprehensively in the use of Thai EFL learners. Three of them are single-word discourse markers: *like*, *so* and *well*. The other three are multi-word discourse markers: *you know*, *I think* and *I mean*. Previous research examined these six spoken discourse markers and discovered that they were all used frequently by native English speakers (Aijmer, 1997; Brinton, 2008, 2010; Tree, 2010). Hence, this research examined how Thai EFL learners used these six spoken discourse markers and made a comparison to the use of them by native English speakers.

This research aimed at investigating both the frequency and the pragmatic functions of the six spoken discourse markers used by Thai EFL learners. First, it aimed at examining whether there was a significant difference in the use of the spoken discourse markers between Thai EFL learners and native English speakers. Second, it examined whether there was a significant difference in the use of the spoken discourse markers between Thai B1-level EFL learners and Thai C1-level EFL learners. Third, it identified the pragmatic functions of the spoken discourse markers used by Thai EFL learners. Fourth, it identified the different pragmatic functions of the spoken discourse markers used by Thai EFL learners compared to native English speakers.

Based on the objectives of the research, four research questions were proposed as follows.

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

Question 2: Is there a significant difference in the use of spoken discourse markers between Thai B1- level EFL learners and Thai C1-level EFL learners?

Question 3: What are the pragmatic functions of spoken discourse markers used by Thai EFL learners?

Question 4: What are the different pragmatic functions of spoken discourse markers used by Thai EFL learners compared to native English speakers?

Based on the research questions, four hypotheses were proposed as follows.

Hypothesis 1: There is a significant difference in the frequency in the use of spoken discourse markers between Thai EFL learners (including both Thai B1-level EFL learners and Thai C1-level EFL learners) and native English speakers.

Hypothesis 2: There is a significant difference in the frequency in the use of spoken discourse markers between Thai B1-level EFL learners and Thai C1-level EFL learners. Generally speaking, Thai B1-level EFL learners use spoken discourse markers at a lower frequency compared to Thai C1-level EFL learners.

Hypothesis 3: The number of pragmatic functions found in using spoken discourse markers is in this order: Thai B1-level EFL learners < Thai C1-level EFL learners < native English speakers.

Hypothesis 4: Both-level Thai EFL learners use basic pragmatic functions of spoken discourse markers and mainly tend to stick to certain pragmatic functions repeatedly.

To answer all the research questions, 30 Thai B1-level EFL learners and 30 Thai C1-level EFL learners were involved in this research. Each participant had an approximately 20-minute English conversation with the researcher on the selected topics such as university life, friends, travel experience, and so forth. Each conversation was recorded and transcribed into written form. The researcher built the learner corpus TELSEC with its two sub-corpora: BTELSEC and CTELSEC. The researcher retrieved the spoken discourse markers based on Brinton's (2008) framework and the criteria of each spoken discourse marker in this research.

In the following parts of this chapter, it presents the main findings of this research, the implications of this research, the limitations of this research and suggestions for future studies.

5.1 Main findings of the research

Since this research examined the spoken discourse markers from two dimensions: the frequency and the pragmatic functions in comparisons between Thai EFL learners and native English speakers, as well as the comparisons between Thai B1-level EFL

learners and Thai C1-level EFL learners, in this part, it presents the main findings of the two comparisons, respectively.

5.1.1 Comparisons of using the spoken discourse markers by Thai EFL learners and native English speakers

Overall, the results from question 1 illustrated the fact that Thai EFL learners used the spoken discourse markers not at similar frequencies as the use of them by native English speakers. It was found that the general underuse of the spoken discourse markers by Thai EFL learners was attributable to three manifestations as follows.

First, Thai EFL learners underused the majority of the pragmatic functions of the spoken discourse markers investigated in this research (59% of the pragmatic functions were underused). Aligned with similar results of previous research (Aijmer, 1997, 2001, 2004; Polat, 2011; Tagliamonte, 2005; Tree, 2010), it exhibits at least two important facts that need further consideration. For one thing, it might be considered that the underuse of the spoken discourse markers is a phenomenon for non-native English speakers in which they produce them less frequently in English conversation than native English speakers. It is assumed, as cited by previous research (Asik & Cephe, 2013; Fung & Carter, 2007; Polat, 2011), that the input of the spoken discourse markers for non-native English speakers is unsystematic and limited. For another, given that it was believed that the importance of the spoken discourse markers came from their contribution to the pragmatic functions in conversation (Bolden, 2015; Brinton, 2008), wherein they signaled the relationship between the speaker and the hearer in the different contextual environments and marked boundaries between two utterances in conversation, it is assumed that it is not an easy task for EFL learners to perform spoken discourse markers to achieve pragmatic competence in communication. It is hence associated with the English proficiency of EFL learners as has been illustrated in the previous part. In other words, a requirement of a higher level of Thai EFL learners' English proficiency may be needed for the improvement in the use of the spoken discourse markers in communication.

In the meantime, as the detailed data and the analysis discussed in the previous part, Thai EFL learners generally performed fewer pragmatic functions than native

English speakers (Thai EFL learners performed 70% of the pragmatic functions compared to native English speakers). They tended to focus on the pragmatic functions that were derived from the original semantic meanings of the spoken discourse markers.

Moreover, as has been illustrated in the answer to question 4, only a small number of Thai EFL learners used each spoken discourse marker in conversation. From these three manifestations, it is therefore concluded that Thai EFL learners generally underused the English spoken discourse markers in conversation.

In addition, as the data shows, it is considered that Thai EFL learners had more deficiency in performing the interpersonal functions of the spoken discourse markers in oral communication. Two interpersonal-centered spoken discourse markers should be paid attention: the discourse marker *well* and the discourse marker *you know*. This is because the majority of the pragmatic functions of *well* and *you know* were underused by Thai EFL learners; 14 out of 27 pragmatic functions of both spoken discourse markers were excluded from the data in TELSEC; and only 6 out of 60 participants used both spoken discourse markers in TELSEC.

It is worth pointing out the phenomenon that the discourse marker *I think* was the only one that was overused by Thai EFL learners in this research. Both levels of Thai EFL learners tended to rely on using the discourse marker *I think* to mark their personal opinion, feeling or stance. They seldom used other discourse markers for the same pragmatic function. Hence, it is believed that Thai EFL learners over-relied on the discourse marker *I think* in its pragmatic function: an implicature of personal opinion, feeling or stance.

Last but not least, four functions need Thai EFL learners' attention due to no or extremely low performance shared by various spoken discourse markers compared to native English speakers. Two of them are textual functions: searching for the following words and as a repair, and the other two are interpersonal functions: as a mitigator and marking opinion, feeling or stance. As has been discussed in the previous part, if Thai EFL learners perform both textual functions above in certain situations, the redundancy of using fillers can be avoided to make the conversation more consistent and coherent. In the meantime, when a potential argument or tension

between the speaker and the hearer appears in conversation, the use of certain spoken discourse markers may ease the tension and face-threat problem. For the function: marking opinion, feeling or stance, Thai EFL learners may improve to using different spoken discourse markers rather than only focusing on the spoken discourse marker *I think*.

Based on all the results shown above, it leads to the fact that Thai EFL learners lack pragmatic competence in oral communication in terms of the discrepancy in the performance of the spoken discourse markers compared to native English speakers. It reflects that Thai EFL learners may be exposed to an unnatural linguistic input in a traditional environment where the focus of English learning lies in grammar and propositional meanings of vocabulary because of the low frequency in the use of the spoken discourse markers in general.

5.1.2 Comparisons of using the spoken discourse markers between Thai B1-level EFL learners and Thai C1-level EFL learners

As illustrated by the answer to question 2, it was found that there was a significant difference in using the spoken discourse markers between Thai B1-level EFL learners and Thai C1-level EFL learners. Upon the examination of each spoken discourse marker, one discourse marker, *you know*, did not have a significant difference, indicating that there was no significant improvement of the performance in *you know* in English conversation with the improvement of the English proficiency from B1-level to C1-level, whereas Thai EFL learners made a significant improvement of the performance in the other five spoken discourse markers in conversation. Overall, it can be concluded that with the improvement of English proficiency, the performance of the spoken discourse markers increased proportionally.

It should be noted that it was found that only a small number of Thai EFL learners, both B1 level and C1 level, performed the spoken discourse markers, especially the cases of the discourse marker *well* and *you know*, wherein only three Thai C1-level EFL learners performed *well* and five Thai EFL learners performed *you know*. Although there was a significant improvement in using the spoken discourse markers from B1-level to C1-level, as illustrated by the findings of question 1, the

276

performance of the spoken discourse markers by Thai C1-level EFL learners still had a significant difference from the performance of native English speakers.

The largest significant improvement was the performance of the spoken discourse marker *like* from B1-level to C1-level wherein Thai C1-level EFL learners performed it much more frequently than Thai B1-level EFL learners, indicating that with the improvement of English proficiency from B1-level to C1-level, Thai EFL learners tended to rely on using *like* as a discourse marker in English conversation.

On the contrary, another point concerning the performance of *like* is worth mentioning. It was found that Thai EFL learners, especially Thai C1-level EFL learners, performed *like* as a filler without any specific semantic or pragmatic meaning between two utterances. Three cases were found in BTELSEC, while 30 cases were found in CTELSEC. This phenomenon, as cited by Aijmer (2004), is the redundancy of some small word between two utterances in conversation, at the moment when non-native English speakers fill the blanks in that they are thinking of the next utterance. There is no agreement on the point whether too many fillers in conversation were considered a bad phenomenon, but the redundancy of too many fillers is considered a special phenomenon for non-native English speakers, which shows the lack of the ability to use English in communication (Aijmer, 2004). It is hence concerned by this research that the performance of the redundancy of *like* should be worth noticing, and avoidance of this phenomenon may seem reasonable for coherent communication.

5.2 Implications of the research

Based on the main findings of the research, this part presents the theoretical implications and pedagogical implications.

5.2.1 Theoretical implications

This research discusses the theoretical implications from two points based on the results of this research: different frequencies and different pragmatic functions of the spoken discourse markers used between Thai EFL learners and native English speakers.

The result of the general underuse of the spoken discourse markers by both Thai B1-level EFL learners and Thai C1-level EFL learners, aligned with several previous research (Aijmer, 2011; Buysse, 2012; Diskin, 2017; Polat, 2011), illustrated that non-native English speakers lacked the use of the spoken discourse markers in English conversation, indicating a deficiency in pragmatic competence in interaction. This finding leads to three theoretical implications.

The first theoretical implication is that EFL learners or non-native English speakers have a deficiency in the awareness of using English spoken discourse markers in oral communication for pragmatic purposes. For one thing, instead of using the spoken discourse markers, EFL learners or non-native English speakers tend to use more fillers, such as *emm* or *uhmm*, to fill the space between utterances. For another, EFL learners or non-native English speakers do not think of using spoken discourse markers for achieving interpersonal interactions in English conversation.

The second theoretical implication is that a low performance of English spoken discourse markers reveals a deficiency in pragmatic competence and pragmatic fossilization in oral communication. In line with some previous research (Asik & Cephe, 2013, Trillo, 2002), a lack of using the spoken discourse markers in English conversations was suggested as a lack of pragmatic competence by EFL learners, in which a competent English speaker should have the ability to express the correct semantic meaning and interacting with other interlocutors at the pragmatic level. Furthermore, both a lack of using some spoken discourse markers or an over-reliance on some spoken discourse markers were considered as inappropriateness of using English spoken discourse markers, resulting in pragmatic fossilization by EFL learners or non-native English speakers (Trillo, 2002).

Another interesting theoretical implication is that the improvement of using the English spoken discourse markers increases proportionally with the improvement of English proficiency of EFL learners or non-native English speakers. As was also illustrated in previous research (Fung & Carter, 2007), it was found that the higher the English level was, the more the spoken discourse markers were used. It indicates that it may be considered to improve the English proficiency to enhance the use of the spoken discourse markers at the same time.

5.2.2 Pedagogical implications

Based on the findings of this research and the theoretical implications, five pedagogical implications are presented.

First, it is appropriate for EFL learners to acquire the concept of the spoken discourse markers in communication in the first place, including the definition of this category, their features, clear clarification of the written discourse markers and the spoken discourse markers, etc. Some discourse markers appear in both written form and spoken form, such as *so*, *I think*, etc. Some discourse markers are only or mainly used in one of the forms, for instance, the discourse marker *well* is considered to appear in the spoken form only. EFL learners or non-native English speakers should be guided to understand the differences in both types of discourse markers. All the knowledge related to the concept of the discourse markers must be introduced as clearly as possible to EFL learners or non-native English speakers so that they are able to establish a complete systematic foundation of this whole category for future studies.

Second, guidance on the importance of the spoken discourse markers in communication should be involved. Based on the previous research (Aijmer, 2004), the spoken discourse markers appear in a high frequency used by native English speakers in daily conversation, and they are able to provide different kinds of pragmatic purposes in different contextual environments. It is hence essential for EFL learners to understand the fact that the spoken discourse markers not only connect two utterances, but also provide different types of pragmatic meanings in communication.

Under this circumstance, EFL learners should have a systematic input and exposure to the main pragmatic functions of each spoken discourse marker. Brinton's (2008) framework is a good point to start the introduction of the functions of the spoken discourse markers because the framework constructs all types of detailed functions of different discourse markers to guide learners for a systematic input. It is worth mentioning that the guidance of the framework is considered to be significant for EFL learners. This is because similar to the definition of the discourse markers, the pragmatic functions of the spoken discourse markers were found in a variety by different researchers so that EFL learners may easily get confused if no framework is

279

introduced at this point. With the guidance of the framework, EFL learners can start to study the functions of each discourse marker based on the findings of the previous research. It is assumed that the original data from the native English speaker corpus can be used as authentic materials for EFL learners to study.

Meanwhile, four pragmatic functions, including two textual functions and two interpersonal functions, should be guided more to Thai EFL learners. For the two textual functions, it is suggested that instead of using fillers such as *emm*, *uhhm*, etc., EFL learners may use different spoken discourse markers in different contexts to have the signal for searching for the right words; and they can perform various discourse markers for a self-correction in the function as a repair. For the two interpersonal functions, on the one hand, the function, as a mitigator, should be introduced to EFL learners in terms of the meaning of this function and when EFL learners can use it in conversation; on the other hand, the function, marking personal opinion, feeling or stance, should be guided to more Thai EFL learners for the purpose of the expansion its use by Thai EFL learners.

Last but not least, based on the result of this research and some previous research (Nookam, 2010), Thai EFL learners should pay attention to the spoken discourse marker *well* and *you know* because these two discourse markers were used the least from the perspectives of their frequencies and functions. For Thai B1-level EFL learners, they should have more guidance on noticing the spoken discourse marker *well* and its functions. For Thai C1-level EFL learners, they should be guided to acquire more different functions that both spoken discourse markers can provide. Meanwhile, Thai EFL learners also need to avoid the over-reliance on the discourse marker *I think* in oral communication. Suggested by previous research (Dehe & Wichmann, 2010), they can be guided to use similar spoken discourse markers, such as *I believe*, *I suppose*, etc. Thai EFL learners should understand the point that not all clauses that express personal opinion, feeling or stance need to start with the collocational pattern *I think*.

5.3 Limitations of the research

The limitations of this research are as follows.

First, this study focused on six spoken discourse markers that have been mostly investigated in the previous research, which were acknowledged as the spoken discourse markers used by native English speakers in high frequencies in English conversation with various pragmatic functions (Aijmer, 1997, 2011, 2014; Bolden, 2015; Brinton, 1996, 2008, 2010). However, this research did not study other English spoken discourse markers. This is because, for one thing, there is no agreement or inventory on this whole category; for another, there still exists a controversy on whether some words or short items can be treated as spoken discourse markers.

Second, this research adopted SBCSAE as the native English speaker corpus mainly because it is comparable and practical to be used in this research, i.e., it is the same genre as the learner corpus built in this research in daily English conversation covering different areas in the USA; and it has free access to the public. As illustrated in the previous chapter, different findings may and most likely will occur should comparisons be made with other English dialects in another native English speaker corpus.

Last but not least, as also illustrated in the previous chapter, all the function analysis was completed by the researcher alone in accordance with Brinton's (2008) framework, conclusions of each pragmatic function of each spoken discourse marker extracted from previous research and Relevance Theory. Although additional explanation has been stated in the functions that may exist ambiguity, a different interpretation may occur should another researcher do the analysis.

5.4 Recommendations for further research

The recommendations for further research are proposed based on the methodology and the results of this research.

First, as suggested earlier, future studies should have a clear classification of the type of discourse markers they are intended to investigate. In this research, the discourse markers referred to the spoken discourse markers that served the pragmatic functions in English daily conversation. But as has been discussed in the previous chapters, the concept of the discourse markers is dependent on different researchers' perspectives. It is thus significant for future studies to make it clear in the definition of

the discourse markers and the type of the discourse markers in the first place. Otherwise, the research may be confusing to the readers.

Second, as mentioned in the limitations of this research, only six English spoken discourse markers were studied because they appeared in high frequencies used by native English speakers and their pragmatic functions played important roles in oral communication (Aijmer, 1997, 2004; Brinton, 1996, 2008, 2010; Tree, 2010). For further research, more discourse markers that are used by Thai EFL learners are suggested to study. Based on the design of some previous research, future research can study individual discourse marker used by non-native English speakers. For instance, Aijmer (2016) studied the discourse marker *anyway*, Haselow (2011) studied the discourse marker *then*, Schourup (2011) studied the discourse marker *now*, etc. Another way of studying discourse markers is that the research scrutinizes a frequency list of all the discourse markers used by non-native English speakers. For example, Asik and Cephe (2013) scrutinized a list of the discourse markers used by Turkish EFL learners; Fung and Carter (2007) had a list of discourse markers used by Hong Kong EFL learners, etc.

Third, this research used SBCSAE as the native English speaker corpus for the main reason that it was a representative of spoken American English in the genre of informal daily English conversation. Hence, SBASAE was comparable to the learner corpus built in this research. However, for future research, other native English speaker corpus in other English dialects is encouraged to use for more comparisons. If other native English speaker corpus is used in further research, the researchers can investigate more different results from the comparisons between non-native English speakers and native English speakers.

Fourth, considering the study of the functions of certain discourse markers, the pragmatic functions of the discourse markers used by native English speakers should be studied further. For one thing, although some discourse markers have been studied several times, their pragmatic functions have not been in total agreement, or new functions may be found due to the language change. For another, the functions of some discourse markers have not been fully studied. The discourse markers, such as *now* (Schourup, 2011), *no* (Lee-Goldman, 2010), *then* (Haselow, 2011), have not been

studied in different corpora to have a further investigation of their pragmatic functions used by native English speakers. Hence, it is worth continuing the study of the pragmatic functions of different discourse markers by using the data of native English speakers.

Last but not least, the further research should continue the study of different discourse markers used by non-native English speakers. On the one hand, studies of the frequency in the use of different discourse markers by non-native English speakers can help the researchers and learners to realize whether they underuse or overuse certain discourse markers. On the other hand, studies of the pragmatic functions of different discourse markers used by non-native English speakers can reveal what functions non-native English speakers use in high frequencies, and what functions non-native English speakers seldom use or do not use at all. Further studies on the use of different discourse markers by non-native English speakers in different backgrounds can make a landscape of how non-native English speakers produce the discourse markers. It may help EFL learners to improve their awareness of using the English discourse markers in both written form and spoken form.

REFERENCES

Adams, C. (2002). Practitioner review: The assessment of language pragmatics. *Journal of Child Psychology and Psychiatry*, *43*, 973-987.

Alghamdi, E. A. (2014). Discourse markers in ESL personal narrative and argumentative papers: A qualitative and quantitative analysis. *International Journal of Humanities and Social Science*, *4*(4), 294-305.

Allen, J. F., Schubert, L., Ferguson, G., & Heeman, P. (1995). The trains project: A case study in building a conversational planning agent. *Journal of Experimental and Theoretical AI*, *7*, 7-48.

Anderson, G. (2000). The role of the pragmatic marker like in utterance interpretation. In G. a. F. Andersen, T (Ed.), *Pragmatic markers and propositional attitude* (pp. 17-38). John Benjamins.

Anderson, G. (2001). *Pragmatic markers and sociolinguistic variation: A relevance theoretic approach to the language of adolescents*. John Benjamins.

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.

Asr, F. T., & Demberg, V. (2013). On the information conveyed by discourse markers. *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics*, 84-93.

Aull, L. (2019). Linguistic markers of stance and genre in upper-level student writing. *Written Communication*, *36*(2), 267-295.

Aull, L. L., & Lancaster, Z. (2014). Linguistic markers of stance in early and advanced academic writing: A corpus-based comparison. *Written Communication*, *31*(2), 151-183.

Babanoglu, P. (2014). A corpus-based study on the use of pragmatic markers as speech-

like features in Turkish EFL learners' argumentative essays. *Procedia, Social and Behavioral Sciences*, *136*, 186-193.

Bax, S., Nakatsuhara, F., & Waller, D. (2019). Researching L2 writers' use of metadiscourse markers at intermediate and advanced levels. *System*, *83*, 79-95.

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

Benveniste, E. (1966). *Problémes de linguistique générale*. Gallimard.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *The Longman grammar of spoken and written English*. Longman.

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.

Blakemore, D. (1987). *Semantic Constraints on Relevance*. Blackwell.

Blakemore, D. (1992). *Understanding utterances. An introduction to pragmatics*. Blackwell, Cambridge MA & Oxford.

Blakemore, D., & Gallai, F. (2014). Discourse markers in free indirect style and interpreting. *Journal of Pragmatics*(60), 106-120.

Blass, R. (1990). *Relevance relations in discourse: A study with special reference to Sissala*. Cambridge University Press.

Bolden, G. (2006). Little words that matter: Discourse markers so and "oh" and the doing of other-attentiveness in social interaction. *Journal of Communication*, *56*, 661-688.

Bolden, G. (2009). Implementing incipient actions: The discourse marker so in English conversation. *Journal of Pragmatics*, *41*, 974-998.

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

Borderia, S. P. (2006). A functional approach to the study of discourse markers. *Approaches to Pragmatic Particles*, 1-46.

Brinton, L. (1996). *Pragmatic markers in English*. Mouton de Gruyter.

Buysse, L. (2012). So as a multifunctional discourse marker in native and learner speech. *Journal of Pragmatics*, *44*, 1764-1782.

Callies, M. (2015). Using learner corpora in language testing and assessment: Current practice and future challenges. In E. Castello, Acerley, K. & Coccetta, F (Ed.), *Studies in learner corpus linguistics: Research and applications for foreign language teaching and assessment* (pp. 21-35). Peter Lang.

Carter, R. A., & McCarthy, M. J. (2006). *Cambridge grammar of English: A comprehensive guide to spoken and written grammar and usage*. Cambridge University Press.

Chase, B. T., & Johannsen, K. L. (2012). *Pathways 2*. Heinle.

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].

Cieri, C., Graff, D., Kimball, O., Miller, D., & Walker, K. (2005). *Fisher English training parts 1 and 2, Speech and transcripts. Linguistic data*. Consortium.

Cieri, C., Miller, D., & Walker, K. (2004). The Fisher corpus: a resource for the next generations of speech-to-text. Proceedings of LREC 2004: Fourth International Conference on Language Resources and Evaluation, Lisbon.

Day, R., & Yamanaka, J. (2003). *Impact topics*. Longman Asia EIT.

Dehe, N., & Wichmann, A. (2010). Sentence-initial I think (that) and I believe (that): Prosodic evidence for use as main clause, comment clause and discourse marker. *Studies in Language*, *34*(1), 36-74.

Diessel, H., & Tomasello, M. (2001). The acquisition of finite complement clauses in English: A corpus-based analysis. *Cognitive Linguistics*, *12*(2), 97-141.

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.

Dobrovoljc, K. (2017). Multi-word discourse markers and their corpus-driven identification. *International Journal of Corpus Linguistics*, *22*(4), 551-582.

Du Bois, J. W., Chafe, W. L., Meyer, C., A, T. S., Englebretson, R., & Martey, N. (2000-2005). *Santa Barbara corpus of spoken American English, Parts 1-4*. Linguistic Data Consortium.

Eales, F., & Oakes, S. (2011). *Speak out*. Pearson Longman.

Erman, B. (1987). *Pragmatic expressions in English*. Almqvist & Wiksell.

Erman, B. (2001). Pragmatic markers revisited with a focus on you know in adult and adolescent talk. *Journal of Pragmatics*, *33*, 1337-1359.

Europe, C. o. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Press Syndicate of the University of Cambridge.

Fitzmaurice, S. (2004). Subjectivity, intersubjectivity and the historical construction of interlocutor stance: from stance markers to discourse markers. *Discourse Studies*, *6*(4), 427-448.

Fox, T., & Schrock, J. C. (2002). Basic meanings of you know and I mean. *Journal of Pragmatics*, *34*, 727-747.

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

Fraser, B. (1990). An approach to discourse markers. *Journal of Pragmatics*, *14*, 383-395.

Fraser, B. (1993). Discourse markers across language. *Pragmatics and language Learning*, *4*, 1-16.

Fraser, B. (1996). Pragmatic markers. *Pragmatics*, *6*(2), 167-190.

Fraser, B. (1999). What are discourse markers? . *Journals of Pragmatics*, *31*, 931-952.

Fraser, B. (2006). Towards a theory of discourse markers. *Approaches to discourse particles*, 189-204.

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

Fraser, B. (2015). The combing of discourse markers: A beginning. *Journal of Pragmatics*, *86*, 48-53.

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

Gilquin, G., Granger, S., & Paquot, M. (2007). Learner corpora: The missing link in EAP pedagogy. *Journal of English for Academic Purposes*, *6*, 319-335.

Granger, S. (1998a). The computer learner corpus: A versatile new source of data for SLA research. In S. Granger (Ed.), *Learner English on Computer* (pp. 3-18).

Addison Wesley Longman.

Granger, S. (Ed.). (1998b). *Learner English on computer*. Addison Wesley Longman.

Granger, S. (2003). Error-tagged learner corpora and CALL: A promising synergy. *CALICO Journal*, *20*(3), 465-480.

Granger, S. (2004). Computer learner corpus research: Current status and future prospects. *Language and Computers*, 123-145.

Granger, S. (2013). A bird's eye view of learner corpus research. In S. Granger, J. Hung, & S. Peter-Tyson (Eds.), *Computer learner corpora, Second language acquisition and foreign language teaching* (pp. 3-33). Benjamins.

Gut, U. (2012). The LeaP corpus: A multilingual corpus of spoken learner German and learner English. In T. Schmidt & K. Wörner (Eds.), *Multilingual Corpora and Multilingual Corpus Analysis* (pp. 3-23). Benjamins.

Halliday, M. A. K., & Hasan, R. H. (1976). *Cohesion in English*. Longman.

Hardie, A. *Welcome to the UCREL significant test system*. Retrieved March 1 from http://corpora.lancs.ac.uk/sigtest/

Heeman, P. A., Byron, D., & Allen, J. F. (1998). Identifying discourse markers in spoken dialog. *Technical Report*, 44-51.

Hellermann, J., & Vergun, A. (2007). Language which is not taught: The discourse marker use of beginning adult learners of English. *Journal of Pragmatics*, 157-179.

Hernandez, T., & Rodriguez-Gonzalez, E. (2013). Impact of instruction on the use of L2 discourse markers. *Journal of Second Language Teaching and Research*, *2*(1), 3-32.

Hiranras, J. (2011). Grammaticalization of the Verb Ɂaw into Pragmatic Markers in Thai. *Journal of Letters*, *40*(1), 100-131.

House, J. (2009). Subjectivity in English as Lingua Franca discourse: The case of you know. *Intercultural Pragmatics*, *6*(2), 171-193.

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.

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.

Janin, A. (2003). *The ICSI meeting corpus* ICASSP, Hong Kong.

Janin, A. (2004). *The ICSI meeting project: Resources and research* NIST ICASSP 2004 Meeting Recognition Workshop, Montreal.

Jendryczka-Wierszycka, J. (2009, July). Collecting spoken learner data: Challenges and benefits. A Polish L1 perspective. Proceedings of the Corpus Linguistics Conference, UK.

Jeon, B., & Perderson, M. (2009). *Dynamic listening & speaking 2*. Compass Publishing.

Jones, C., & Carter, R. (2012). Teaching spoken discourse markers explicitly: A comparison of III and PPP. *International Journal of English Studies*, *14*(1), 37-54.

Jucker, A. H. (1993). The discourse marker well: A relevance theoretical account. *Journal of Pragmatics, 19*, 435-452.

Jucker, A. H., & Ziv, Y. (1998). Discourse markers: Introduction. In A. H. Jucker & Y. Ziv (Eds.), *Discourse markers: Descriptions and theory* (pp. 1-12). John Benjamins Publishing Company.

Khatib, M., & Safari, M. (2011). Comprehension of discourse markers and reading comprehension. *English Language Teaching*, *4*(3), 243-250.

Khedri, M., Heng, C. S., & Ebrahimi, S. F. (2013). An exploration of interactive metadiscourse markers in academic research article abstracts in two disciplines. *Discourse Studies*, *15*(3), 319-331.

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

Lam, P. (2009). Discourse particles in corpus data and textbooks: The case of well. *Applied Linguistics*, *31*(2), 260-281.

Lee-Goldman, R. (2010). No as a discourse marker. *Journal of Pragmatics*, *43*, 2627-2649.

Leech, G., Garside, R., & Bryant, M. (1994). CLAWS4: The tagging of the British National Corpus. Proceedings of the 15th International Conference on

Computational Linguistics (COLING 94), Kyoto, Japan.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Lawrence Erlbaum Associates.

Meurers, D. (2012). Learner corpora and natural language processing. *Semantic Scholar*, 1-27.

Modhish, A. S. (2012). Use of discourse markers in the composition writing of Arab EFL learners. *English Language Teaching*, *5*(5), 56-61.

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

Muşlu, M. (2018). Use of stance lexical bundles by Turkish and Japanese EFL learners and native English speakers in academic writing. *Gaziantep University Journal of Social Sciences*, *17*(4), 1319-1336.

Myles, F. (2005). Interlanguage corpora and second language acquisition research. *Second Language Research*, *21*(4), 373-391.

Neff, J. A., Ballesteros, F., Dafouz, E., & Martinez, F. (2004). Formulating Writer Stance: A Contrastive Study of EFL Learner Corpora. *Languages and Computers*, 562-571.

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]. Thailand.

Norrick, N. R. (2001). Discourse markers in oral narrative. *Journal of Pragmatics*, *33*, 849-878.

Pasaribu, T. A. (2017). Male and female students' use of textual discourse markers in writing academic essays. *Journal of Language and Literature*, *17*(1), 74-81.

Pelteret, C. (2012). *Speaking B1+Intermediate*. HarperCollins Publishers.

Polat, B. (2011). Investigating acquisition of discourse markers through a developmental learner corpora. *Journal of Pragmatics*, *43*, 3745-3756.

Popescu-Belis, A., & Zufferey, S. (2010). Automatic identification of discourse markers in dialogues: An in-depth study of like and well. *Computer Speech and Language*, *25*, 499-518.

Povolna, R. (2012). Casual and contrastive discourse markers in novice academic

writing. *Brno Studies in English*, *38*(2), 131-148.

Pravec, N. A. (2002). Survey of learner corpora. *ICAME Journal*, *26*, 81-114.

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.

Rahimi, M. (2011). Discourse markers in argumentative and expository writing of Iranian EFL learners. *World Journal of English Language*, *1*(2), 68-78.

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.

Rayson, P., & Garside, R. (2000, 1-8 October). Comparing corpora using frequency profiling. workshop on Comparing Corpora, held in conjunction with the 38th annual meeting of the Association for Computational Linguistics (ACL 2000), Hong Kong.

Redeker, G. (1990). Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics*, *14*, 367-381.

Rezanova, Z. I., & Kogut, S. V. (2015). Types of discourse markers: Their ethnocultural diversity in scientific text. *Procedia: Social and Behavioral Sciences*, *215*, 266-272.

Ross, S. M. (2017). *Introductory statistics*. Academic Press.

Sadeghi, B., & Heidaryan, H. (2012). The effect of teaching pragmatic discourse markers on EFL learners' listening comprehension. *English Linguistics Research*, *1*(2), 165-176.

Sakita, T. I. (2013). Discourse markers as stance markers. *Pragmatics Cognition*, *21*(1), 81-116.

Schegloff, E. A. (1992). Repair after next turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology*, *97*(5), 1295-1345.

Schegloff, E. A. (2001). Getting serious: Joke serious no. *Journal of Pragmatics*, *33*,

1945-1955.

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

Schourup, L. (2001). Rethinking well. *Journal of Pragmatics*, *33*, 1025-1060.

Schourup, L. (2011). The discourse marker now: A relevance-theoretic approach. *Journal of Pragmatics*, *43*, 2110-2129.

Simeikaite, A. (2012). Spoken discourse markers in learner academic writing. *Studies about Language*, *20*, 27-33.

Simma, P. (2014). Function of SENG 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). In

Tagliamonte, S. (2005). So who? Like how? Just what? Discourse markers in the conversations of Young Canadians. *Journal of Pragmatics*, *37*, 1896-1915.

Tantiwich, K., & Sinwongsuwat, K. (2019). Thai university students' use of yes/no tokens in spoken interaction. *English Language Teaching*, *12*(3), 1-14.

Tree, J., E, F. (2010). Discourse markers across speakers and settings. *Language and linguistic compass*, *4*(5), 269-281.

Vanda, K. H., & Peter, F. B. (2011). Gender differences in the use of the discourse markers you know and I mean. *Argumentum*, *7*, 1-18.

Vande-Kopple, W. (1985). Some exploratory discourse on metadiscourse. *College Composition and Communication*, *36*, 82-93.

Waltereit, R., & Detges, U. (2007). Different functions, different histories. Modal particles and discourse markers from a diachronic point of view. *Catalan Journal of Linguistics*, *6*, 61-80.

Watts, R. J. (1988). A relevance-theoretic approach to commentary pragmatic markers: The case of actually, really and basically. *Acta Linguistica Hungarica*, *38*, 235-260.

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.

Yang, S. (2011). Investigating discourse markers in pedagogical settings: A literature review. *ARECLES*, *8*, 95-108.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Appendix

Sample questions in the conversations with Thai EFL learners

Instructions:

1. Each participant was told that their personal information would not be given in any circumstance, including their real names, age, gender, faculty, etc.

2. Each participant was told that all the conversations would be recorded in audio.

3. Each participant was told that certain content of the conversation would appear in the research paper without any personal information to the public.

4. Each participant was told that the conversation should be an informal daily conversation.

Questions relevant to the topic of "education":

1. Tell me something about your school life before university. What was it like?

2. What was your senior high school life like?

3. What do you think of your university life so far? How is it different from your school life?

4. What do you think of your teachers, the campus or anything else related to your university?

5. What do you want to improve for the rest of your university life?

Questions relevant to the topic of "people":

1. How do you describe your personality?

2. Have you kept in touch with your friends in primary school or high school?

3. How do you usually make friends?

4. What do you usually do with your friends? Is anything special that you can think of?

5. Are you feeling comfortable now? What makes you uncomfortable?

Questions relevant to the topic of "problems and solutions":

1. Was there anything that bothered you when you were growing up?

2. How did you solve the problem when you were bothered by the things you just mentioned when you were growing up?

3. Has anything bothered you in your university life so far?

4. Do you have any problems living in Bangkok?

5. How do you usually solve the problems that occur during the study?

Questions relevant to the topic of "family":

1. What was your childhood like?

2. Do you think the relationship between you and your parents is very close?

3. What is the best thing about your family?

4. Did you have any arguments with your family?

5. If you have your own family in the future, what do you want to do to make your family close?

Questions relevant to the topic of "travel":

1. Do you like traveling? Why?

2. Tell me one place that you visited before.

3. What do you usually do during traveling?

4. Are you willing to share any special experiences during a trip before?

5. Which country do you want to visit in the future? Why?

VITA

NAME	Zhaoyi Pan
DATE OF BIRTH	10 January 1989
PLACE OF BIRTH	Nanjing, Jiangsu Province, China
INSTITUTIONS ATTENDED	English as an International language, Chulalongkorn University
HOME ADDRESS	2099 Ideo Mobi Sukhumvit Road Phra Khanong Bangkok 10260 Thailand