rEFLections
Vol 31, No 2, May - August 2024

Differentiating Learner's English Proficiency in Oral Presentation Focusing on Textual Markers

SASITHORN LIMGOMOLVILAS
Chulalongkorn University Language Institute, Chulalongkorn University, Thailand
Author email: sasithorn.li@chula.ac.th

Article information

Article history:
Received: 17 Mar 2024
Accepted: 26 Jul 2024
Available online: 7 Aug 2024

Keywords:
Assessment
Coherence
Oral presentation
Textual metadiscourse
Transitional markers

Abstract

Creating its own learner corpus, this research proposed to analyze and classify the transition markers used in solo presentation by 30 Thai engineering students based on Hyland (2019)'s Marker Categorization in Textual Metadiscourse. This research also aimed to identify and compare the quantities of individual transition markers among three groups of the selected 30 students (10 high, 10 mid, and 10 low users). The research employed both quantitative and qualitative method through SPSS, Transkriptor, and AntConc. The results showed that -and- was the top word utlized as a textual marker among all three groups while some other top transition markers ranked slightly differently among the groups. In addition to the search on the list of textual markers, the manual observation encountered other phrases acting as textual markers. The results indicated that users from average level can apply phrases as textual markers to communicate coherence instead of using traditional transition markers. Moreover, the "reminders", the act of reminding the listeners of previous content, manifested as a possible indicator of advanced proficiency users as the weak users were found to not produce any of this category. Thus, learners may improve their English-speaking skills if they are guided through the process of developing coherence in context using various categories of textual markers.

INTRODUCTION

Various learner corpora have been developed around the world, both in English and other languages, such as Arabic, Chinese, Italian and Spanish (Centre for English Corpus Linguistics, 2023). Many learner corpora that focus on the general English language use have been formed (Granger, 1998; Meyer 2002). This availability includes sources for English Learner's corpora that are publicly available for studies, such as CORFL or CORpus of English as a Foreign Language. Learner corpora can be employed to improve teaching strategies (Meyer, 2002) by studying various aspects involved such as influence of native language and error patterns. In language testing, learner corpora can be utilized at many levels, such as defining user needs, defining test purpose, and determining the learner's capability or proficiency level (Barker, 2010). In Thailand, most learner corpora have been developed for small scale private research, but there is one corpus available for public access: the Thai Learner English Corpus (TLEC hereafter), which was created in 2009 by the Department of Linguistics, Chulalongkorn University.

TLEC is composed of mainly essays ranging from intermediate, advanced, and professional level, but originated solely from written corpus.

Learner corpora can reveal the learners' pattern of language use at diverse levels (Barker, 2010). It is agreed by many researchers (e.g., Barker, 2010; Flowerdew, 2012) that an important advantage of learner corpora is the capability in detecting and analyzing learners' errors. Granger (1998) added that error detection will be extremely useful when conducted on a specific group of learners. Not only defining common language-use patterns, but a learner corpus can also be used for defining the common errors that the learners make. Both can be the detail for the descriptors listed in the rubric. Thus, finding errors and classifying them from learner corpus can help increase awareness of learner's level and act as a possible influence on how to succeed at a higher level (Crosthwaite & Raquel, 2019). In other words, the information on the common pattern language use and learners' error can be indicators and examples of performance for each level. To make such justifications, a comparison of learner's language use is required, which is a comparison between native speakers of English and the group of English learners.

Some studies (e.g., Bui, 2022; Chotimongkol et al., 2011; Thumawongsa, 2018) have aimed to compare language use between non-native and native speakers of English. For instance, Chotimongkol et al. (2011) analyzed the phonetic errors of Thai learners while reading using a contrastive list. The comparison of learner's data to the native speakers of English may seem risky, but it can inform us as to what difficulties these learners are facing (Granger & Tribble, 2014). Such recognition can be beneficial to both learners and instructors. For example, a study (Thumawongsa, 2018) on Thai learners' preposition usage found that some errors stemmed from the interference of the mother tongue. In such a case, students can be alerted to the specific prepositions since they may use the direct translation approach to communicate. Another study reflecting L1 interference was reported by Bui (2022) on the cohesive use of Vietnamese students, who used the translation method and this resulted in the wrong meaning and structure of several words, such as *so* and *therefore*.

Studies (e.g., Arya, 2020; Pan & Aroonmanakun, 2022) examining the linguistic features of the writing ability of Thai students through topics such as conjunction, intensifier, part of speech, relative marker and tense can be found. However, there is not much research involving the use of textual markers in Thai learner spoken corpus. Some studies (e.g., Arya, 2020; Mamuenvai & Rhekhalilit, 2021) focused on a few simple keywords, which may not be used for much of the interpretation. Furthermore, the nature of grammar used in writing differs from that used in speaking (Jones & Waller, 2015). A research study concerning a wider list of the textual markers used in the speaking of Thai learners may enhance the perspective of assessment of this relatively scarce spoken corpus. In addition, revealing the common pattern of the group may help the learners to understand and avoid the mistakes as well as help the instructors to understand and tackle the problem wisely. This study uses the data from a group of Thai students, whose English language is assumed to be learned as a foreign language. This study aims to analyze the different characteristics of the language use of learners in three proficiency levels according to their performance in oral presentation.

LITERATURE REVIEW

Various issues can be considered in assessing a student's performance. In this study, language use is the focus in dividing students' level. According to Bachman (1990), language knowledge is categorized into two primary domains: organizational knowledge and pragmatic knowledge. These two domains can be interpreted as a steppingstone of language capability use starting with organizational knowledge as the first step that language users need to be equipped with followed by pragmatic knowledge at a more advanced level to use language to match purposes or situations. Organizational knowledge is the fundamental building block of language formation, which is the use of vocabulary, syntax, and phonological or graphological information. The textual knowledge continues after the basic components in grammatical knowledge are grouped to form an ordered text. Textual knowledge encompasses knowledge of rhetorical or conversational organization, which helps the language user arrange the text in ways such as comparison-contrast texts and turn-taking in conversations, as well as knowledge of cohesiveness, which joins sentences with linking words.

Table 1

Crismore et al.'s marker categorization in textual metadiscourse (Adapted from Hyland, 2019, p. 40)

Markers	Category	Function	Examples
Textual	Logical connectives	Show connections between ideas	therefore; so; in addition; and
	Sequencers	Indicate sequence/ ordering of material	first; next; finally; 1, 2, 3
	Reminders	Refer to earlier text material	as we saw in Chapter one
	Topicalizers	Indicate a shift in topic	well; now I will discuss …
Interpretive	Code glosses	Explain text material	for example; that is
	Illocution markers	Name the act performed	to conclude; in sum; I predict
	Announcements	Announce upcoming material	In the next section

In discourse analysis, metadiscourse is an approach to understand textual knowledge and the persons who use it (Hyland, 2019). One way how metadiscourse is applied widely is to set an explicit textual device as the focal point (Hyland, 2019). Even though metadiscourse markers were established in written context, Hyland (2019) endorses metadiscourse "as central to the overall purpose of language use" (Hyland, 2019, p. 42). Hyland (2019) then referred to Vande Kopple's system of categorizing metadiscourse into seven features but mentioned that the narration and attributes of these features can be hard to differentiate as they can overlap and complement each other's function. However, an improved version of Vande Kopple's system was proposed by Crismore et al. (1993), which Hyland (2019) summarized as the changes of textual metadiscourse, as shown in Table 1. Crismore et al. (1993)'s textual metadiscourse is divided into two main categories of textual markers and interpretive markers. The textual markers can be categorized through logical connectives, sequencers, reminders and topicalizers, while the interpretive markers are classified as code glosses, illocution markers and announcements. Although these categorizations can be a clear medium to identify the words used for metadiscourse, transition markers seem to capture both concepts of textual and interpretive markers. As termed by Hyland (2019), transition markers can be used to demonstrate how each idea is connected through addition, comparison and consequence.

The application of textual markers can be an indicator of a better English user. Learners who master both linguistic and textual knowledge in terms of cohesiveness can produce texts that

are not only grammatically correct but also clear, coherent, and engaging for the reader (Bachman, 1990). Many corpus-based studies have been conducted with their focus on how cohesive devices can differentiate the learner's proficiency level, especially in their writing. For instance, Yang and Sun (2012) assert that cohesion is a necessary tool for assessing the quality of writing as well as specifying it as a characteristic of learners with advanced proficiency level. Similarly, Bax et al. (2019) investigated the use of metadiscourse markers among three levels, which were B2, C1, C2, and found that learners tend to increase the overall use of metadiscourse markers to certain levels that they have learned in order to adapt and apply a more advanced method to make the text coherent and minimize the number of explicit markers. In addition, the results indicated that a wider variation of discourse markers is applied in learners at high proficiency level, which is aligned with the study (Bui, 2022) revealing the belief of Vietnamese students that the more cohesive devices used the more advanced they are perceived to be. In other words, it is believed that more proficient learners have a wider variety of language knowledge, mainly of word choice and of how to express those words, to be applied in their written text. However, Chung et al. (2023) could not indicate any pattern on the textual markers among the three levels and even reported that the overuse of transition markers does not lead to a higher score.

In terms of cohesion in an oral context, applying textual devices may yield a resemblant result and some practical recommendations. Although Crosthwaite and Raquel (2019)'s studies are on various linguistic features applied in group discussion, they found that the use of discourse markers can be a significant predictor to a higher score to sustain fluency during a talk, especially the markers for engagement and topic shift. As a matter of fact, McCarthy (2010) pinpointed that rather than grammatical errors, it is the fluency and confluence of the talk that is the main pivot for the rater in judging a learner's performance. Nonetheless, Pan and Aroonmanakun (2022) observed that when compared to native speakers, Thai EFL is deficient in applying some spoken discourse markers for pragmatic purposes. This comparison does not address Krashen's input theory in applying it to different proficiency levels. Assembling more corpus learners from different levels for research can help to determine confluent factors for effective communication (McCarthy, 2010). Exploring the application of textual markers among Thai learners from different levels of proficiency can also fill in such gaps.

In this study, students were taught to construct their presentations with an introduction, a body, and an ending as well as to apply transition markers. To complete the list of transition markers in the presentation, considering only logical connectives and sequencers cannot complete the circuit of presentation as they do not include the ending or the conclusion. Since this study aimed to analyze a 3–4-minute solo presentation by students in three proficiency levels and to be precise in scope, investigating the use of textual markers under the four categorizations mentioned above seemed to best fit the context of this analysis.

Objectives

1. To analyze and classify the transition markers used in solo presentation by Thai students.
2. To identify and compare the quantities of individual transition markers among three groups – proficient users, average users, and weak users.

METHODOLOGY

This study employed secondary data source from an assessment of solo presentations, which was one of the tasks in a listening and speaking course for all engineering students at a university in Thailand. The students from all majors are required to take this course as the third English requirement out of four courses; the students take this course during different semesters depending on their major. The arrangment of such registration for the course is organized by the faculty. As a result, some engineering students would study this course as early as the first semester of their second year while others would take it as late as the penultimate semester in their fourth year for the four-year Bachelor's program. In this course, students are prepped to enter the work field, and areas covered include job application, job interview, group discussion, and solo presentation. The task required students to perform a solo presentation on any topic that was engineering related with a length of 3 to 4 minutes. Normally, students had to perform face-to-face in the classroom with 2 raters and 4 students as the audience. However, due to the COVID situation, students were allowed to submit a video record online three weeks after the assessment was announced. A guideline and rubric were posted and explained to the students in class with some practice and consultation with the class instructor. This research aimed to analyze the use of textual markers of Thai students at three levels (proficient, average, and weak) and consisted of the following steps: data collection, data transcription and data analysis.

Solo presentation task

Solo presentation task is the last task in the course for students to complete and submit to receive 15% of the total score. The task requires students to present any technical content that is related to engieneering within 3–4 minutes. As mentioned earlier, all students studying in the Thai engineering program are required to study this course: however, the content in engineering can vary depending upon the student's major and own particular interest. Students were prepared to perform the task for about one month. Then, the selected topic and content had to be preapproved by the class instructor. Students prepared the content to practice with their instructor who provided comments for students to improve their performance in presentation skills.

A rubric (See Appendix) on the solo presentation was provided and explained to students as well. The rubric ranges from superior level to failure in three criteria, which are content, delivery, and visuals. The content includes story and organization, which mentions the use of organization as "flawless use of varied transition and explanatory language" and "organization of the talk is not only intelligent but also interesting/clever".

Since the classes were held during the COVID pandemic, students completed their solo presentation in video form and were not allowed to edit the audio components of their video; however, they were allowed to edit the visual elements to incorporate their slides. A link to Google Drive per class was provided for students to upload the presentation and the presentation slides. After that, the students' presentations were assessed independently by two raters, neither of whom was their class instructor. The scores and comments were entered on Google

sheet. A discussion session was needed when the difference in score between the two raters was over 3 points out of 30.

Data collection

Not all of the videos submitted could be used for analysis as the performance of some of the students was unacceptable in terms of task fulfillment. A lot of the videos turned in were described as reading scripts as students only displayed the script on screen while recording the presentation. As a result, the first step in data collection was to eliminate those videos with comments from the raters indicating reading from the script as that would affect the real speaking performance. The scores of 30 video files were selected and divided into three groups: 10 proficient, 10 average and 10 weak proficiency levels. Although the full score consists of three parts (content, delivery, visuals), only the delivery scores were considered. Based on a full score of 10, the data was categorized into 3 groups of proficient, average, and weak group according to the score scale of 9–10, 7–8, and 6–5, respectively.

Data transcription

Transkriptor, an online program, was utilized to transcribe the videos from spoken to written text. Since the program is AI assisted and may miss some words that the students mispronounced or may correct the grammar for students who make mistakes, the researcher would have to listen and check the words transcribed from the program before saving the script for analysis in a text file format. First, phonetic problems were detected during the transcribing process. As an illustration, the program can analyze the words students pronounced wrongly and can offer a word that matches the context, and that did happen in some cases. If the speakers had a heavy Thai accent, the program or the researcher could not find a matching word as the original words were virtually unintelligible. Listening to transcribe the audio only would not produce the right word; however, watching the visual could sometimes help in replacing the missing words. A second problem stemmed from the degree of perfection and correctness that the program perceives based on the language of a native speaker of English. With regard to the grammar mistakes that the speakers made, the program corrected the grammar in the script, which does not reflect the real language used by the speaker. Having the researcher recheck the original script helped in solving these two main problems.

Data analysis

Both qualitative and quantitative analysis were employed through SPSS version 19 and AntConc 4.2.4. To prove pairwise differences among the three groups, the delivery scores of the 30 participants were analyzed with F-test (One-Way ANOVA) tests and the Scheffé method for multiple comparisons. For qualitative analysis, AntConc 4.2.4 (MacOS 10/11-Silcon) was applied to analyze a list of transition markers individually. Since the list of the transition words was too short, the transition markers had to be entered and recorded word by word. Two categories are “logical connectives” and “sequencers”, which are derived from a list of a computer tool called Text Inspector (Bax, 2012), as well as label stages since it is a list of words that is under the category of transition markers. Another two, which are “reminders” and “topicalizers”, had to be defined manually as there is no specific word list to match.

RESULTS

The results from ANOVA show that the three groups are significantly different in terms of delivery score, as is shown in Table 2. In addition, comparisons between each pair of the three groups were tested to confirm the difference with the Scheffé method. Table 3 illustrates that the three groups are significantly different from each other based on the scores.

Table 2

ANOVA results on delivery score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	42.650	2	21.325	88.241	.000
Within Groups	6.525	27	.242		
Total	49.175	29			

Table 3

Multiple comparisons using the scheffé method

Conditions	Mean Difference	Std. Error	Sig.	95% Confidence Interval	
				Lower Bound	Upper Bound
Proficient Vs. Average	1.523*	.201	.000	1.001	2.044
Proficient Vs. Weak	3.107*	.229	.000	2.512	3.701
Average Vs. Weak	1.584*	.233	.000	.980	2.188

The mean difference is significant at the 0.05 level.

Divided into three proficiency levels, the corpus in this study consists of 13,967 tokens. After accumulating the corpus, the first objective was to analyze and classify the transition markers used in solo presentation by Thai students, while the second aim was to identify and compare the quantities of individual transition markers. Based on the results, all students applied textual devices in their talk as each was a prepared presentation following on from a consultation with their instructor. Responding to both objectives, Table 4 summarizes the top seven transition markers with the range use of over 16 out of 30 participants as well as the frequency and the percentage compared to the entire list of transition markers and to the whole data. The words at the top of the list of transition markers of this learner corpus are: *and*, *or*, *so*, *but*, *first*, *because*, and *second*. Other words were combined in the same category in the last line of the table at 48, 61, and 60 words for the proficient, average, and weak levels respectively.

Table 4

A list of transition markers of all groups with frequency and percentage in learner corpus

Words	Proficient			Average			Weak		
	Frequency	Markers (%)	In Data (%)	Frequency	Markers (%)	In Data (%)	Frequency	Markers (%)	In Data (%)
and	139	45.87	2.73	114	39.31	2.39	110	43.48	2.68
or	30	9.90	0.59	20	6.90	0.42	25	9.88	0.61
so	32	10.56	0.63	23	7.93	0.48	14	5.53	0.34
but	22	7.26	0.43	21	7.24	0.44	16	6.32	0.39

Words	Proficient			Average			Weak		
	Frequency	Markers (%)	In Data (%)	Frequency	Markers (%)	In Data (%)	Frequency	Markers (%)	In Data (%)
first	11	3.63	0.22	21	7.24	0.44	14	5.53	0.34
because	16	5.28	0.31	15	5.17	0.32	5	1.98	0.12
second	5	1.65	0.10	15	5.17	0.32	9	3.56	0.22
Others	48	15.84	0.94	61	21.03	1.28	60	23.72	1.46
Total	303	100	5.95	290	100	6.09	253	100	6.16

Proficient users

A total of 22 different transition markers was detected from the corpus of the proficient users. Other words for proficient users were: *second*, *finally*, *then*, *after*, *let's start*, *and then*, *let's begin*, *moreover*, *last*, *next*, *however*, *therefore*, *in conclusion*, *lastly*, and *firstly*. The word with the highest frequency of use - *and* - was divided into its use as either a conjunction or a transition marker, which was when *and* was used in the initial position of the sentence. It was found that the word *and* was used 80 times to function as a conjunction and 59 times as a transition marker. *And* was found to occur with the word *then* five times.

Table 5

A list of transition markers used by proficient users with frequency and percentage in learner corpus

Transition markers	Frequency	Transition marker (%)	In Data (%)	Range
and	139	45.87	2.73	10
so	32	10.56	0.63	10
or	30	9.90	0.59	9
but	22	7.26	0.43	8
because	16	5.28	0.31	8
also	16	5.28	0.31	6
first	11	3.63	0.22	6
second	5	1.65	0.10	4
finally, then	4	1.32	0.08	3
after, let's start	3	0.99	0.06	3
and then	5	1.65	0.10	2
let's begin, moreover, last, next	2	0.66	0.04	2
however, therefore, in conclusion, lastly, firstly	1	0.33	0.02	1

Average users

A total of 26 transition markers was detected from the corpus of average users as shown in Table 6. Other words for average users were: *after*, *next, let's start*, *then*, *however*, *again*, *also*, *and then*, *last*, *lastly*, *in conclusion*, *after that*, *third*, *since*, *let's begin*, *after all*, *finally*, *and et cetera* and *thus*. When dividing the word with the highest frequency of use - and - into its use as a conjunction and as a transition marker it was found that the word was used 56 times to function as a conjunction and 58 times as a transition marker.

Table 6

A list of transition markers used by average users with frequency and percentage in learner corpus

Transition markers	Frequency	Transition marker (%)	In Data (%)	Range
and	114	39.31	2.39	10
first	21	7.24	0.44	9
but	21	7.24	0.44	9
so	23	7.93	0.48	8
or	20	6.90	0.42	6
second	12	5.17	0.32	6
because	15	5.17	0.32	5
after	8	2.76	0.17	5
next	7	2.41	0.13	5
let's start	6	2.07	0.13	5
then	5	1.72	0.11	5
however, again	4	1.37	0.08	3
also	3	1.03	0.06	3
and then	3	1.03	0.06	2
last, lastly	3	1.03	0.06	1
in conclusion, after that, third	2	0.69	0.04	2
since, let's begin, after all	2	0.69	0.04	1
finally, and et cetera, thus	1	0.34	0.02	1

Weak users

A total of 26 transition markers was detected from the corpus of weak users. Other words for weak users are *next*, *last*, *also*, *finally*, *let's start*, *third*, *firstly*, *secondly*, *thirdly*, *after that*, *then*, *and then*, *moreover*, *in conclusion*, *let's begin*, *as a result*, *conversely*, *and so on*, and *in addition to*. When the most frequently used word - *and* - was divided into its use as a conjunction and a transition marker, it was found to function 63 times as a conjunction and 47 times as a transition marker. The range of *next* for weak users was not wide. In fact, *next* was found to be at the top list from the overuse of one speaker who used it 10 times while other speakers used it only once.

Table 7

A list of transition markers used by weak users with frequency and percentage in learner corpus

Transition marker	Frequency	Transition marker (%)	In Data (%)	Range
and	110	43.48	2.68	10
or	25	9.88	0.61	7
but	16	6.32	0.39	5
first	14	5.53	0.34	6
so	14	5.53	0.34	7
next	13	5.14	0.32	4
last	9	3.56	0.22	7
second	9	3.56	0.22	6
also	6	2.37	0.15	5
because	5	1.98	0.12	4
finally	4	1.58	0.10	2
let's start	3	1.19	0.07	3
third	3	1.19	0.07	2
firstly	5	1.98	0.12	2
secondly	2	0.79	0.05	2
thirdly, after that	2	0.79	0.05	1
then	3	1.19	0.07	1
and then, moreover, in conclusion, let's begin, as a result, conversely, and so on, in addition to	1	0.40	0.02	1

As mentioned under Data Analysis, there was no list for "reminders" and "topicalizers" to match like "logical connectives" or "sequencers", so these lists had to be specified through manual reading. Another 18 transition markers were found to be utilized as shown in Table 8, where the results are divided into 4 categories. The word *well* was used most often by both the proficient and the weak users. The proficient users were found to employ other transition markers the most frequently, especially the reminders. The average users came second and tended to utilize more of the sequencers. The weak users were the group with the highest use being of topicalizers.

Table 8

A list of other transition markers and frequency

Categories	Other markers	Proficient	Average	Weak
Logical connectives	What's more…	-	-	1
	Another serious problem is …	-	1	-
Sequencers	Before we talk about…,	-	1	-
	Before you know about that … / Before we know how…	-	2	-
	Before we dive into …	-	1	-
	Before we get into the detail of …, let's talk about…	1	-	-
	Before I finish, let me …	2	1	-
Reminders	So now you already know about …	1	1	-
	After we know about …, let's talk about…	-	1	-
	Now you know that…	1	-	-
	Now that I've talked about…	1	-	-
	Like I mentioned before, ...	1	-	-
	Back to the example that I showed you.	1	-	-
	This is related to what I said earlier.	1	-	-
Topicalizers	Now you may have a question what is …	1	-	-
	Well, …	4	-	3
	You might think that is good? Should I change it? Don't worry, I will show you …	-	-	1
Total		14	8	5

DISCUSSION

As this assessment allowed students to present within the time limit of no more than four minutes, not much technical content could be included. Based on the use of sequencing words, the highest number of subtopics that could be covered seemed to be three. This significantly affected the number of transition markers that could be applied. In other words, the time limitation and the requirement that the topic had to be explained to a non-technical audience could have influenced the number of transition markers used.

Since *and* is one of the top five most frequent function words (Meyer, 2023), it is not surprising to see that 'and' is the top textual marker found in this study. Generally speaking, the function of *and* is not limited to that of transition marker, but often *and* is used as a conjunction. As the results showed, around 49-57% of the use of *and* was found to be as a conjunction. It should be noted that extent of *and* used was about the same for all proficiency levels, which concurs with other studies (e.g., Arya, 2020; Mamuenvai & Rhekhalilit, 2021). This indicates

that using *and* as a transition marker does not affect the scores in this learner corpora, especially when the aim of this speaking task was to maintain fluency, naturalness, and enthusiasm, but not variety and accuracy. In fact, *and* is a very fundamental word for EFL learners to use for addition (Arya, 2020).

Considering other discourse markers that are not on the list, two that are worth mentioning are *well* and *so*. The word *well* was found to be used the most frequently, but not as frequently as is used by native speakers to maintain flow of interaction (McCarthy, 2010). The results in this study support a previous study (Pan & Aroonmanakun, 2022) on how Thai learners tend to underutilize *well* in a spoken context. Nevertheless, the use of *so* in this study contradicts what Pan and Aroonmanakun (2022) believed to be one of the underused oral discourse markers of Thai learners. This could derive from the different context of oral communication of the corpus as the context of this corpus is a solo presentation on technical content that may require more use of transition markers than everyday conversational dialogue.

The different variety of transition markers used among the three groups is intriguing. It is plausible that some transition markers were viewed by the proficient students as inappropriate for spoken context. Walková (2020) explained this situation as stylistic awareness, which is how some transition markers are believed to be proper for some contexts but not for others. Similarly, other studies (Chung et al., 2023; Khaghaninejad et al., 2021) have found that the number of sequence markers applied varies according to genre. In this speaking, a genre in which students were asked to present the knowledge of their field to a non-technical audience, most of the transition markers used were simple as the speaker was trying to convey the meaning of technical words and terms to laymen audiences.

Another gripping reason for the lesser number of transition markers of proficient users might be the avoidance of some formal transition markers. As the results revealed that reminders were mainly used by profient users and from one average user but none from the weak users, this suggests that reminders are possibly an indicator of advanced proficiency users. In other words, applying unfixed transition markers or phrases to communicate coherence is an ability performed by proficient and average users. This follows what Bax et al. (2019) suggested: that when the learners become more proficient, they can organize a better written text without depending on the routinely used markers, which is also consistent with what Chung et al. (2023) found. As the label suggests, the reminders were applied to remind and conclude each subtopic for the listeners. Making certain that the listener can follow the idea and the technical content may contribute as a factor in being a good communicator. As a matter of fact, the act of applying their own phrases as a transitional marker is described in the rubric of superior level as “flawless use of varied transition and explanatory language”.

A couple of pedagogies from the findings of the textual markers in this learner corpora can be implied. First, not all transition markers suit an oral presentation. Considering the audience’s understanding of techincal content and context may overcome the need to apply some fancy transition markers. In fact, instructors should suggest that learners try to keep their organization and content as clear as possible for the audience as well as to apply simple transitional markers. Next, the use of a reminder as an act of referring to previous materials should be stressed to

the learners as it can encourage the coherence of the content. The act can be conducted with simple phrases which should not be too difficult for the low learners to apply and to enhance the coherence in content. Nevertheless, instructors should not advise learners to always use the same phrases as it can result in fixed transitional markers like other textual markers found in logical connectives and sequencers. Last but not least, although the results suggest a difference in the textual markers among each level, they do not indicate that textual markers as function words are the main indicators for determinging either the learners' proficiency level or the specific field of a student's major.

LIMITATIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

Focusing on the textual markers use of Thai engineering students, it can be hard to generalize the results to other EFLs. Apart from the subject matter of this learner corpus, the scope of textual markers can also be a limitation of this study. To cover all the markers in this context, inclusion of interpretive markers might offer different points of view. In addition, this corpus collected information from only Thai engineering students; understanding would be enhanced if similar assessment was applied to more diverse groups of Thai learners. A larger size of learner corpus with various ranges of speakers could enlarge the generalizability of the results.

CONCLUSION

This study analyzed the transition markers, identified and compared the quantities used in solo presentation by Thai students at three levels: proficient, average, and weak. The scores from the three groups were analyzed to confirm that they were significantly different from each other. The number of the transition markers was found to be equivalent among all groups. Similar to what had been found in previous studies, *and* was the top word utlized as a textual marker even after classsifying from conjunction. The rank of the other top seven words was slightly different among the three groups. The data in the corpus suggested that most of the speakers followed the same structure since everyone received guidance on constructing a presentation by starting with an opening, followed by a body and a conclusion as well as using transition markers to guide the audience. However, the transition markers applied by the proficient users were found to be plain but specific to the context. This study suggests that learners can appear coherent by applying simple words or phrases to connect the ideas and maintain their fluency. Using phrases to function as a reminder is considered a pluasible factor to differentiate students as advanced level. Although plenty of fixed sequencers were avaliable, some average users produced a phrase to avoid using the usual transition markers. Weak users tended to use fixed transition markers as recommended and could not generate their own phrase as a transition marker, neither sequencers nor reminders. Facilitating the learners through these stepping stones of creating coherence in context through other categories of textual markers may help learners to become better English speakers during a presentation.

THE AUTHOR

Sasithorn Limgomolvilas is a full-time lecturer at Chulalongkorn University Language Institute (CULI). She graduated from Chulalongkorn University with a Bachelor's degree in Education. After receiving her Master's degree in Teaching English as a Speaker of Other Languages (TESOL) from San Francisco State University, she returned to Thailand to teach at CULI before graduating with a Ph.D. in English as an International Language (EIL) from Chulalongkorn University. Her main research interests are in ESP and language assessment.
sasithorn.li@chula.ac.th

REFERENCES

Arya, T. (2020). Exploring discourse marker use in Thai university students' conversations. *LEARN Journal: Language Education and Acquisition Research Network, 13*(1), 247–267. https://so04.tci-thaijo.org/index.php/LEARN/article/view/237854

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.

Barker, F. (2010). How can corpora be used in language testing? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics*. Routledge.

Bax, S. (2012). *Text inspector* [Online text analysis tool]. https://textinspector.com/

Bax, S., Nakatsuhara, F., & Waller, D. (2019). Researching L2 writers' use of metadiscourse markers at intermediate and advanced levels. *System, 83*, 79–95. https://doi.org/10.1016/j.system.2019.02.010

Bui, H. P. (2022). Vietnamese EFL students' use and misconceptions of cohesive devices in writing. *SAGE Open, 12*(3), Article 1126993. https://doi.org/10.1177/21582440221126993

Chotimongkol, A., Thatphithakkul, S., Chootrakool, P., Hansakunbuntheung C., & Wutiwiwatchai, C. (2011). The design and development of PELECAN: Pronunciation errors from learners of English corpus and annotation. *International Conference on Speech Database and Assessments, 2011*, 36–41. https://doi.org/10.1109/ICSDA.2011.6085976

Chung, E., Crosthwaite, P., & Lee, C. (2023). The use of metadiscourse by secondary-level Chinese learners of English in examination scripts: insights from a corpus-based study. *International Review of Applied Linguistics in Language Teaching, 62*(2), 977–1008. https://doi.org/10.1515/iral-2022-0155

Crosthwaite, P. R., & Raquel, M. (2019). Validating an L2 academic group oral assessment: Insights from a spoken learner corpus. *Language Assessment Quarterly, 16*(1), 39-63. https://doi.org/10.1080/15434303.2019.1572149

Flowerdew, L. (2012). *Corpora and language education*. Palgrave Macmillan.

Granger, S. (1998). The computer learner corpus: A testbed for electronic EFL tools. In J. A. Nerbonne (Ed.), *Linguistic databases* (pp. 175–188). CSLI Publications.

Granger, S., & Tribble, C. (2014). Learner corpus data in the foriegn language classroom: Form-focused instruction and data driven learning. In S. Granger & G. Leech (Eds.), *Learner English on computer* (pp. 199–209). Routledge.

Hyland, K. (2013). Metadiscourse. In K. Hyland (Ed.). *Discourse studies reader: Essential excerpts* (pp. 65–88). Bloomsbury Academic. http://dx.doi.org/10.5040/9781472541925.ch-004

Hyland, K. (2019). *Metadiscourse: Exploring interaction in writing*. Bloomsbury Academic.

Jones, C., & Waller, D. (2015). *Corpus linguistics for grammar: A guide for research* (1st ed.). Routledge. https://doi.org/10.4324/9781315713779

Khaghaninejad, M. S., Eslami, M., Yadollahi, S., & Jafari, S. M. (2021). A corpus based analysis of the application of "concluding transition signals" in academic texts. *Cogent Arts & Humanities, 8*(1), Article 1868223. https://doi.org/10.1080/23311983.2020.1868223

Mamuenvai, P., & Rhekhalilit, K. (2021). A corpus-based study of AND by Thai EFL learners. *The Proceedings of the 59th Kasetsart University Annual Conference, 2021*, 355–363.

McCarthy, M. (2010). Spoken fluency revisited. *English Profile Journal, 1*, Article e4. https://doi.org/10.1017/S2041536210000012

Meyer, C. F. (2002). *English corpus linguistics: An introduction*. Cambridge University Press.

Meyer, C. F. (2023). *English corpus linguistics: An introduction* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/9781107298026

Pan, Z., & Aroonmanakun, W. (2022). A corpus-based study on the use of spoken discourse markers by Thai EFL learners. *LEARN Journal: Language Education and Acquisition Research Network, 15*(2), 187–213. https://so04.tci-thaijo.org/index.php/LEARN/article/view/259927

Thumawongsa, N. (2018). L1 transfer on Thai EFL learners' utilization of prepositions: A corpus-based analysis. *Journal of Community Development Research (Humanities and Social Sciences), 11*(1), 35–47.

Walková, M. (2020). Transition markers in EAP textbooks. *Journal of English for Academic Purposes, 46*, Article 100874. https://doi.org/10.1016/j.jeap.2020.100874

Yang, W., & Sun, Y. (2012). The use of cohesive devices in argumentative writing by Chinese EFL learners at different proficiency levels. *Linguistics and Education, 23*(1), 31–48. https://doi.org/10.1016/j.linged.2011.09.004

Appendix

Content (story & organization)

Fail	Poor	Average	Strong	Superior
1-3	4-5	6-7	8-9	10
Well below expected standard; Serious problems with topic/content. Seriously disorganized/unclear. No introduction or conclusion.	- Conveys little to no meaningful information (no substance); audience wonders "what is this *about*?" - Technical details add to confusion rather than clarify - Unclear how (if at all) topic relates to audience concerns. - Is essentially a sales presentation or advertisement OR is essentially a lecture, more suited to a uni course than a general interest presentation - Very little formal intro (perhaps ONLY name and topic) - Seems very disorganized OR illogically organized - No conclusion or very flat ("That is all for my presentation…")	- Conveys information but some key points are unclear - Attempts to clarify technical details for non-specialist audience but with limited success - Attempts to relate topic to audience concerns but vaguely or in an unconvincing way - Uses a hook but not very effectively, e.g. after topic is already announced (anticlimactic) OR overly simplistic "Have you ever wondered about X? Today I will tell you about X) - Attempts to use transitions and internal explanatory language but with mixed results - Presentation may seem to "skip around" with one part somewhat out of sequence. - Attempts conclusion but perfunctory/weak.	- Conveys solid information with all key points clear and well explained - Successfully anticipates needs of a non-specialist audience and clarifies technical details accordingly (analogies, definition of key terms, examples) - Explains to audience how the topic is important to them/their lives. - Uses hook effectively (to create interest followed by topic reveal) - Consistent use of varied transition language and internal explanatory language with few if any lapses - Gives all essential parts of intro [greeting, name, topic, outline, time, question policy] - Has a strong conclusion (clear end to talk, summary, call to action or takeaway, audience questions)	- Highest standard; genuinely educational - Makes something complex comprehensible to non-specialists; - Seems professional quality (TED-worthy) / Is genuinely entertaining. - Hook is genuinely creative and interesting - Flawless use of varied transition and explanatory language - Organization of talk is not only intelligent but also interesting/clever - Conclusion meets all basic requirements and goes beyond: Inspiring or memorable conclusion

Delivery (body language & language)

Fail	Poor	Average	Strong	Superior
1-3	4-5	6-7	8-9	10
Well below expected standard; distraction from eye contact, posture, room position, gestures, or note reading detract seriously from presentation. Genuinely cannot understand student, lapses into Thai, does not speak	- Little eye contact. Turns back to audience at length - Few gestures/hands in pockets or arms crossed/clutching notes - Posture extremely stiff OR excessive/distracting motion - Reads from notes excessively - Pronunciation causes serious difficulty in understanding; listener genuinely does not understand some key points - Monotone - Grammatical issues cause confusion - Extremely choppy/halting delivery, long pauses - Sounds poorly memorized (almost unlistenable)	- Some eye contact but inconsistent - Occasionally turns to or walks in front of slides - Limited use of gestures but inconsistent - Posture weak, seems nervous or fidgets - Use of notes is distracting at times - Pronunciation causes some difficulty in understanding; listener must strain to comprehend at times - Somewhat flat/monotone. Inconsistent use of stress - Pauses too much, OR not often enough, leading to confusing "run on" - Misuses vocab/grammar to the extent that it is distracting or makes it difficult to understand key points - Somewhat halting delivery - Sounds scripted/memorized	- Good eye contact, few if any lapses, spread around room. - Faces audience and doesn't walk in front of own slides, good position in room with few if any lapses - Uses gestures effectively to convey sense of what is being said and call attention to slides - Posture relaxed and confident, nothing distracting - Uses notes sparingly or not at all - Pronunciation is understandable; only very rarely causes difficulty understanding. - Good use of vocal variety. Does not sound scripted. - Uses a level of language appropriate to a formal talk (vocab and grammar) with only minor slips - Reasonably fluent with minor lapses	- Highest standard - Perfect eye contact, genuinely seems to be interacting with audience - Perfect use of space, no need to turn back or even away from audience, knows what is on slides without having to look - "Owns the room" – comes across as completely poised and professional, confident and in control - No use of notes – feels spontaneous or conversational - Pronunciation causes no difficulty in understanding, even if pronunciation is not native. Excellent use of vocal variety - Completely fluent. Sounds natural. - Uses a wide range of vocabulary and grammar accurately

Visuals

Fail	Poor	Average	Strong	Superior
1-3	4-5	6-7	8-9	10
Well below expected standard; No visuals or visuals that seem completely perfunctory, full of errors, inappropriate or disrespectful content; uses video to do work of presentation	- Frequent errors/typos - Blocks of text/confusing or visually overwhelming slides - Data presented unclearly, doesn't help audience to understand point, inappropriate for a talk (e.g. textbook graph with tiny labels) - Low quality or inappropriate (unprofessional) images - Many slides simply have text that's read by speaker - Gratuitous or ineffective use of video: Does not work and is abandoned, intrusive/distracting (opens separate file), overly long with no student input ("let's all watch TV together").	- Several errors/typos - Some slides may be wordy/have extraneous detail - Data not entirely clear, may use graphics or diagrams that contain excessive notations - Slides mismatched (cartoons mixed with high-res photos, Google branded graphic with something found on internet) - Slides may repeat what is being said somewhat, contribute to audience understanding with limited success. - May use video but awkwardly, without full control of playback, or in a way that is not totally necessary to presentation (a still image would have sufficed); allows video narration to do the work that they should be doing	- No errors - Only 1 or 2 mildly wordy / "noisy" slides - Most data simplified and optimized for presentation reasonably well (no copy & paste of a visual that doesn't really work on big screen) - Unified visual style (slides feel like part of same presentation) - Slides actually contribute to our understanding, not merely placeholders or repetition of what's being said - Any use of video well-integrated, fairly well executed, contributes to our understanding and is entirely narrated by the presenter with no distractions	- Highest standard - No errors - Not wordy / Not "noisy" - All data has been simplified and optimized for presentation (no copy & paste of a visual that doesn't really work on big screen) - Single, elegant visual style (style seems consciously picked to fit the topic) - Slides allow the presenter to convey complex information that would be impossible to explain without them. - Any use of video is seamlessly integrated and genuinely contributes something that a still picture couldn't; totally narrated by presenter

Professionalism

Scored by **deduction** out of the total score the student gets. The ideal student is so well prepared that they show up for the assessment on time, look professional, and have so thoroughly prepared and rehearsed their presentation that they can deliver it within the target time (3-4 mins).

Deduct 1 pt: Unkempt / poor self-presentation
Deduct 1 pt: Late to the assessment
Deduct 1 pt: +/-30 seconds OR Deduct 2 pts: +/-30-60 seconds OR Deduct 3 pts: Teacher has to stop student (>5:00)/Short (<2:00)