

ORIGINAL ARTICLE

Open Access



Utilizing an automated tool analysis to evaluate EFL students' writing performances

Anongnad Petchprasert

Correspondence: anongnad07@hotmail.com

Department of Curriculum and Instruction, Faculty of Education, Ramkhamhaeng University, 2086 Ramkhamhaeng Road, Hua Mak, Bangapi Bangkok 10240, Thailand

Abstract

Recently, the integration of linguistics and technology has been promoted and widely used in the field of linguistics and English writing research for several purposes. One of those purposes is to evaluate English as a Foreign Language (EFL) writing ability by using electronic assessment tools. In the current study, an automated writing evaluation tool (Coh-Metrix) was used to indicate English-major students' writing performances based on the discourse components of the texts. The English texts generated for each writing task on two different topics were collected. The corpus analyses gathered from Coh-Metrix identified linguistic and discourse features that were interpreted to determine the 40 EFL undergraduate students' English writing abilities. The students wrote and revised their essays in hand-written essays in class and resubmitted their essays in digital forms with corrections made. The results showed that these students demonstrated linguistic flexibility across writing assignments that they produced. The analyses also indicated that the length of the texts and the uses of the word concreteness, and the referential and deep cohesion had impacts on the students' writing performances across the writing tasks. Besides, the findings suggest practical value in using an automated text analysis to support teachers' instructional decisions that could help to identify improvement of students' writing skill.

Keywords: EFL writing, Coh-Metrix, Discourse components, Corpus linguistics, Computational linguistics

Introduction

The uses and effects of automated tools in analyzing students' writing abilities have been investigated for several years (Buckingham Shum, Sándor, Goldsmith, Bass, & McWilliams, 2017; Haswell, 2000; Ullmann, 2019). However, recent advances in computational linguistics and discourse processing have shown possibilities that educators or researchers can automate many language- and text-processing mechanisms. Several studies have employed automated text analysis methods for educational contexts especially in the area of writing assessment (Ullmann, 2019; Wilson & Czirik, 2016) and discourse analysis (Ferretti & Lewis, 2019).



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

While analyzing text manually poses constraints on writing pedagogy and research, such as, adding time and effort and limiting large-scale research explorations, automated methods tend to come to replace as they have several benefits for writing instruction and research, such as saving time, receiving immediate feedback, and covering multi-language texts (Deane, 2013; Humphreys & Wang, 2017; Ullmann, 2019).

In line with the benefits of implementing automated text analysis, the present study thus investigated the potential of machine language learning algorithms to explain text segments of writing. Also contributing importantly to a discourse model, an automated computational linguistic tool was also used to draw inferences about students' English writing performances.

Linguistics features to analyze argumentative texts

In foreign language learning, students are required to practice their language skills: listening, speaking, reading, and writing. Writing is considered as an important skill in our daily lives and it is definitely one such challenge skill that English as a Foreign Language (EFL) learners must improve.

In academic realm, writing an argumentative text is a way to express a point of view on a particular topic and support it with evidence which is often the aim of academic writing. In particular, EFL writers may be requested to express supported opinions for decision-making, argue for a policy, or evaluate a model developed to solve a particular problem (Zhu, 2001). Therefore, the ability to produce evidence based essays is vital because this type of texts can help individuals persuade, negotiate, debate, or resolve conflicts between people.

In second language (L2) writing research, several studies have examined writing processes and strategies that have been used in generating texts. As an argumentative text has its own complexity, researchers or educators measure students' writing performances in relatively isolated contexts, such as a rhetorical difficulty for second and foreign language writers (Zhu, 2001), an assessment of writing performance (Allen, Likens, & McNamara, 2019), and writing quality (Kim, Gatlin, Otaiba, & Wanzek, 2018). Additionally, some studies have commonly revolved around the assessment and the analysis of the linguistic and rhetorical features of argumentative essays (Ferretti & Graham, 2019; MacArthur, Jennings, & Philippakos, 2019). In most writing research, participants were asked to write essays in response to a single prompt or more prompts. Consequently, they practically performed their abilities to respond to a particular prompt in a relatively controlled environment posing a critical problem. By explaining the problem, the evaluation of high-quality writing is taken into account as they can be varied across raters, authors, assignments, or contexts (Allen, Jacovina, & McNamara, 2016; Varner, Roscoe, & McNamara, 2013). Consequently, this is suggested to adequately examine what components of writing affect students' performances and how students could employ linguistic variation across writing assignments. Several studies have examined writing performance by identifying linguistic features of writing (McNamara, Graesser, McCarthy, & Cai, 2014) involving standardized scoring methods of large essay corpora, and using expert human raters to evaluate quality of essay writing (Deane, 2013).

Crossley, Roscoe, and McNamara (2014) state that high-quality writing has multiple forms to define a successful writer that cannot be identified by a single set of pre-defined linguistic properties. Findings from several studies illustrate the common components of English composition that represent high-quality writing.

For example, Faigley & Witte's study (1981) examined errors and content variables in writing. The results showed that high-rated essays were longer in length and contained fewer errors but more nonrestrictive modifiers. The high-rated essays contained more cohesion than those low-rated essays. Shermis and Burstein (2013) and Deane (2013) have found that linguistic features that commonly distinguish high-rated academic writing are often associated with a greater number of words, especially sophisticated word choices, and a fewer number of spelling and grammar errors. Allen et al., study (2019) has led to an idea that high-quality argumentative writing must contain crucial elements, such as high cohesion and narrative flexibility because these linguistic features have often found associated with essay scores. They also mentioned that research in writing is required to fully comprehend the roles of flexibility in writing processes and writing quality, and examine linguistic flexibility across multiple dimensions in argumentative essay writing.

In accordance with the research and the linguistic features mentioned earlier, the aim of the current study was to examine the discourse components of the argumentative texts analyzed by an automated text tool. The approach used in the study was to consider the notion that there are multiple linguistic dimensions of texts that students generate. Some basic features related to the characteristics of the words, the perceived sophistication of sentences in texts, and the styles of texts can influence readability and complexity of sentence structures (Allen et al., 2016; Allen et al., 2019). In particular, discourse components can be calculated and go beyond the levels of word use and sentence structures in texts. The discourse could also reflect a degree of narrative flexibility in the texts and determine students' writing abilities. In regards to the aim of this study, the descriptions of discourse components and linguistic features implemented as the theoretical framework are presented in Table 1.

Table 1 The six discourse components and linguistic features

Component	Description
Word length	The total number of words found in the text is calculated using the output from the Charniak parser that generates a parse tree with part of speech tags for clauses, phrases, words, or punctuations.
Narrativity	The text tells a story with characters, events, places, and things that are familiar to a reader. Stories are basically about everyday oral conversation.
Syntactic simplicity	This component shows the degree to which the sentences in the text contain fewer or more words with the uses of simple, familiar or complex, unfamiliar syntactic structures.
Word concreteness	Content words are concrete, meaningful, and simple to understand. Abstract words are difficult to represent visually, so the texts that have a lot of abstract words tend to be more challenging than those content words.
Referential cohesion	Higher referential cohesion tends to have words and ideas that extend beyond sentences and the entire text. Lower cohesion is typically more difficult to process or connect the ideas together.
Deep cohesion	This component reflects the degree to which the text contains causal and intentional connectives that help a reader to have a more coherent and deeper understanding of events, processes, or actions in the text.

Note: The description of this framework is adapted from http://141.225.41.245/cohmetrixhome/documentation_indices.html

An automated tool in text analyses

Three main approaches have been introduced to the examination of textual data: the dictionary-based, the rule-based, and the machine learning based approaches. With the following overviews of these approaches, researchers may grasp an idea of each one. The dictionary-based approach does not provide linguistic foundation, nor does it cover the breadth of the method available for examining texts (Tausczik & Pennebaker, 2010). The rule-based approach is implemented to define a variety of rules for getting opinions created by breaking each sentence into a semantic unit for processing. The machine learning based approach refers to a method that uses computers or other automated tools to measure constructs of sentences, provide systematic comparisons, and discover patterns that neither human researchers nor participants of a study can detect. In light of the recent influx of available automated tools for doing writing research, the machine learning based approach has been actively driven by educators and researchers to explicitly demonstrate new ways and patterns to analyze texts (Shermis & Burstein, 2013; Ullmann, 2019).

In terms of text analysis, a question of continuing interests to researchers is what advantages or disadvantages between human coders and computer programs in analyzing texts are. For human coders, Humphreys and Wang (2017) state that an analysis and an interpretation of the results made by human coders are relatively inconsistent since they rely heavily on the expertise and attentiveness of one or more human coders. Besides, conventional measures of the inter-rater reliability for each pair of the human coders are intended to deal with the criteria of replicability (Humphreys & Wang, 2017), meaning that the analysis has to be repeated to ensure the chance of getting the same results. The main notion of replicability is not problematic with an automated text analysis which generates the same results through the repeated processes, thereby contributing validity and reliability. Most importantly, humans may miscode due to unintentional mistakes or biases. After all, researchers may take advantages of computers in an automated process that is good for investigating texts in a context where humans might have limitations. Besides, computers can display patterns in language that humans cannot detect and they can measure textual data consistently and precisely over time through integrating, comparing, and associating with textual data.

Computers or automated text tools can provide information about multiple variables in textual features, explain differential findings of multilingual texts, allow users to easily store results of analyses in data files, and demonstrate cohesion relations and language discourse characteristics (Elfenbein, 2011; McNamara et al., 2014). Besides, an automated text analysis is considered reliable in managing basic to complicated discourse components in linguistics, such as accessing vocabulary from online dictionaries and indicating misspelled words, basic morphemes, or phonemes. Through advances across disciplines has made it possible to conduct computational assessment of language and text comprehension that exceed those basic linguistic elements. The disciplines that advance the field of linguistic studies include computational linguistics (Graesser et al., 2014), analyses of structural and mechanical errors in texts (Crossley, Bradfield, & Bustamante, 2019), and discourse processing (Hardy & Friginal, 2016). By all means, scientific research and language technologies have been integrated, so teachers or researchers can benefit from the automated text analyses. One such level of

language analysis that demonstrates a particular computational challenge is the discourse components of the students' texts.

A variety of available Natural Language Processing (NLP) tools have been introduced and used to demonstrate how these tools evaluate and analyze language discourse of texts. Coh-Metrix is one of those tools that has been widely used in L2 writing research concentrating on writing development, a prediction of quality of texts, an analysis of differences between variations of texts, and an effect of writing tasks. Coh-Metrix has been developed by the researchers from the Institute for Intelligent Systems at the University of Memphis.

As it is aligned within a variety of theoretical framework, Coh-Metrix has been utilized to automatically calculate specific linguistic information ranging from basic text features to higher levels of measurement. Coh-Metrix reports statistical information including the length of specific discourse units within the text, meaning that a report of the analysis will show the number of words and paragraphs, the average length of words, and the number of sentences. At a shallow level of analysis, it is used to calculate the types of words containing within a text, such as the levels of word specificity, the average frequency of individual words, and the diversity of words found in the text. In addition, Coh-Metrix is useful for measuring syntactic complexity; however, it cannot make much sense of larger discourse units such as phases of text. Generally, Coh-Metrix yields the results of analyses on discourse components: number of words, narrativity, syntactic simplicity, word concreteness, referential cohesion, and deep cohesion (see Table 1).

Recently, researchers have also implemented Coh-Metrix to measure multiple levels of English language, such as a textbase, a situation model, and a rhetorical structure. One such effort that has brought to the attention of researchers in writing is a study carried out by Graesser et al. (2014). They reviewed how the five discourse components (narrativity, syntactic simplicity, word concreteness, referential cohesion, and deep cohesion) account for text variations. Varner et al. (2013) implemented Coh-Metrix to examine misalignments between teachers' and students' evaluation criteria for writing quality and to investigate the correlations between the linguistic features and student and teacher ratings of the students' essays. Among others, Aryadoust and Liu (2015) used Coh-Metrix to measure a theoretical model by investigating the associations between the linguistic features of text complexity and text quality. In their study, the differences and similarities in the two sets of linguistic features were investigated. That is, linguistic features including lexical sophistication, syntactic complexity, cohesion, and basic text information were identified to see an association with the writing scores for both integrated and independent sampled groups of Chinese EFL learners. The findings showed that lexical sophistication was a predictor for writing assignments. Besides, syntactic features such as verbs in third-person-singular form and semantic similarity were significantly found as predictors for the integrated writing assignment.

Regarding the literature review on automated tools for text analysis and the theoretical framework used in the present study, Coh-Metrix was used to investigate the components of the EFL students' texts based on the six components: *word length*, *syntactic simplicity*, *word concreteness*, *referential cohesion*, *deep cohesion*, and *number of words* in the students' texts.

Research questions

The current study examined the discourse components of the EFL students' texts to characterize each component and to determine the students' writing performances. The research questions are as follows:

- 1) Based on the results of the text analyses, what are the characteristics of the students' writing performances?
- 2) Is there any significant correlation between the discourse components of the texts across the writing tasks?

In the following sections, a method and an implementation of an automated text evaluation tool are presented in details, starting with a selection of participants and connecting linguistic aspects to important constructs in a text analysis.

Method

Participants

Eighty English-major students were recruited from a large university in Thailand to voluntarily participate in the present study. Eighteen students were screened out prior to beginning the study because they have not completed all prerequisite writing courses, and twenty two students did not have complete writing tasks, giving a final sample size of 40.

Study procedure

During the course of the study, the participants were assigned to write essays in a four-week interval on two different topics over the 8 week period. For the purposes of the study, only two writing tasks (one at the beginning and the last one at the end of the study) were collected to determine the participants' writing performances, investigate the characteristics of the texts, and examine associations between the discourse components of the texts. In doing so, the participants wrote on a total set of two essay topics: 1) *Is education useless in the twenty-first century?* and 2) *A Week Without Access to the Internet*. The participants wrote each essay in a 50 min session and submitted their essays in digital forms as Coh-Metrix only examined the edited digital forms of the texts. The researcher informed the participants that their scores from both tasks would not affect their grades, so they were not allowed to use the automatic grammar and spell checker on the word processor for their texts. In addition, the participants were asked to submit their hand-written drafts and the digital texts to the researcher so that they could be easily reviewed and compared.

Research tool

In the present study, Coh-Metrix (a computational text analysis tool) was used to measure the readability and the complexity of the texts, analyze the texts at the word, sentence, and discourse levels to provide information about the specific linguistic scaffolds containing in the texts across all writing tasks.

Table 2 The descriptive statistics of the number of words in texts

	<i>n</i>	Minimum words	Maximum words	<i>M</i>	<i>SD</i>
Word count on task 1	40	271.00	895.00	511.86	173.00
Word count on task 2	40	296.00	1121.00	575.97	200.63

Results and discussion

As previously stated, the texts were analyzed for *the number of words* and the five discourse components of the students' texts: *narrativity*, *syntactic simplicity*, *word concreteness*, *referential cohesion*, and *deep cohesion* in response to the research questions. The statistical data on the number of words and the other five discourse components are presented in following sections.

Number of words

The number of words contained in the 80 texts is shown in Table 2. The reported descriptive data and the discussions are also presented in the following section.

The average number of words on the writing task 1 was 511 in length, while it was 575 words on the writing task 2. By comparing the number of words from both writing tasks, it was found that 29 participants (69%) wrote longer texts on the second writing assignment, making significantly higher number of words in length, $t(41) = -8.57$, $p < .05$.

Regarding the number of words in texts, findings from several studies suggest that successful writers tend to produce linguistically longer texts (Crossley et al., 2014; McNamara, Crossley, & Roscoe, 2013); however, these results of the analyses cannot solely define a successful writer without investigating further for other linguistic properties or text components that could also influence students' writing performances. Then, further investigations into discourse components and linguistic features were conducted and reported in the following sections.

Discourse components across writing assignments

Research question 1: Based on the results of the text analyses, what are the characteristics of the students' writing performances?

A *z*-score is a standardized metric in standard deviation units, with the value of zero being the mean. The *z*-scores are higher and positive when the texts are easier on the component and more negative when the texts are more difficult. The following sections present statistical data of each discourse component as well as illustrating the results of the analyses in the forms of extracts or passages to characterize participants' writing performances.

Narrativity

Table 3 shows the results of the *narrativity* analysis between the participants' texts generated for the writing tasks. Of the 80 texts, the *z*-scores ranged from -0.4 to -0.5 , making the texts were complicated and rather difficult to read especially on those assigned for the writing task 2. The results also reflect the degree to which a story is being told by using characters, places, events, or other things familiar to readers.

Table 3 The results of the narrativity analysis

z-score range	NT1 (n = 40)	NT2 (n = 40)
-1.0 to < -0.5	3	1
-0.5 to < 0.0	12	16
0.0 to < 0.5	18	13
0.5 to < 1.0	6	7
1.0 to < 1.5	–	2
1.5 to < 2.0	1	1

Note: NT1 = narrativity on writing task one, NT2 = narrativity on writing task two

A passage sample (1) from the first writing assignment illustrates that the writer employed high frequent words and simple syntax in his/her writing. The following passage demonstrates these characteristics.

- (1) Internet is one of the important things to me. *I* spend my time with *it* about 10 hours per day. *I* like to watch the movies or drama series, listen to music, and read some articles in Facebook. Furthermore, *I* often use *it* to communicate with my friends. Nevertheless, *I* cannot connect the internet when *I* come back home.

This passage (1) has a fairly strong beginning as it states how the Internet is important in the writer's life by giving examples in the second and the third sentences. Notice that the writer repeated the pronoun "*I*" several times in almost all the sentences demonstrating that the writer strongly emphasizes personal experiences relevant to the topic. Likewise, Li (2014) states that the use of first person pronoun like '*I*' is identified as a writer's visibility and its over use is regarded as inappropriate and rather informal in academic writing.

The following passage from the writing task 2 exemplifies the different characteristics of the texts through the uses of verbs and intentional actions that the writer used to convey the messages to his/her readers. For example:

- (2) *I use* YouTube to *listen* to music, *watch* movies, and something which *makes* me fun. Accordingly, access to the Internet *makes* me *feel* so relaxed after having studied so hard.

The passage (2) illustrates the uses of *verbs* and *intentional actions* in writer's anticipated experiences of a week without an access to the Internet. The writer used the verbs, such as *use*, *listen*, and *watch* to visualize the actions that he/she routinely does in everyday life.

Syntactic simplicity

Table 4 shows the results of syntactically simple texts between the writing task 1 and 2. In reference to the results, the passage samples are presented to demonstrate the characteristics of the texts.

Table 4 The results of syntactically simple analysis

z-score range	SS1 (n = 40)	SS2 (n = 40)
- 1.5 to < - 1.0	2	1
- 1.0 to < - 0.5	5	8
-0.5 to < 0.0	14	13
0.0 to < 0.5	15	10
0.5 to < 1.0	4	8

Note: SS1 = syntactic simplicity on writing task one, SS2 = syntactic simplicity on writing task two

Statistically, the texts on the writing task 1 ($n = 21$, 52.5%) and on the writing task 2 ($n = 22$, 55%) were less than the mean value, identifying that their levels of syntactic simplicity was not significantly different.

Nevertheless, the researcher investigated in depth found that some participants performed better writing on their second assignment. The following passage samples illustrated this finding. The passages were written by the same student showing the differences of the syntactic structures across his/her writing tasks. The first passage was assigned for the writing task 1 on the topic of “*Is education useless in the twenty-first century?*”

- (3) First of all, learning from schools is really important for children. Schools help their parents for take care of them while they are working. Children can play and meet their friends. It makes them learning each other and learning to handle with people.

With regard to syntax, the writer of the passage (3) tends to write choppy sentences resulting his/her writing relatively unsophisticated and disconnected. The syntactically simple structures apparently demonstrate the participant’s writing ability. In particular, the writer infrequently used transitional words (e.g. *moreover*, *furthermore*, *then*) or coordinating conjunctions (e.g. *and*, *but*, *so*) resulting disconnected ideas in the passage and less expression of the new descriptions to the texts.

Compare the previous passage sample (3) with the following passage (4) from the same writer who wrote it on the topic of “*A Week Without Access to the Internet.*”

- (4) Firstly, Thai students lack of taking their responsibility. In the morning, some students are like to attend the class lately. *Moreover*, they do not realize it will be the reason to lead them to become lazy person. *Furthermore*, some of them are likely to procrastinate on their task or duty. They might look like lazy person. But learning online must take a lot of responsibility on them *because* no one can force or motivate you to do it like learning in class. *Then* this is the reason *that* Thai students are not ready for learning online.

The passage (4) shows the uses of *because* and *that* illustrating a more elaborated text with more complex syntactic structures. This writer also produced a

longer text with less choppy sentences as well as using transitional words (e.g. *moreover, furthermore, then*) more often to describe the associations between the ideas in the passage.

Word concreteness

In general, Coh-Metrix analyzed texts that contain concrete and meaningful words that can easily evoke mental images. A high number of word concreteness corresponds to easier and more understandable texts. The results of the word concreteness analysis are shown in Table 5.

The results of the analysis show that almost all texts across the two writing tasks were less than the mean value (0.0), meaning that the writers tend to use a larger number of abstract words in their writings that generated more complicated texts to read.

The following passages were written by the same writer to illustrate these characteristics:

- (5) Basically, education makes life better. *Educated* people are different from uneducated one in many way, such as *attitude, lifestyle, and social status*. Even at the present time, there are many jobs that everybody can do without knowledge, but those jobs are not good enough to *make a living* and to *be accepted* by society.

By contributing importantly to greater uses of abstract ideas, the writer of the passage (5) on the second writing assignment seems to have in part the abilities to perceive abstract concepts with reference to particular instances (e.g., *educated, attitude, lifestyle*) and verbs (e.g. *make a living, be accepted*) to distinguish relationships among ideas and to express the writer’s experience into the text. This is associated with the writer’s background knowledge and experience he/she shared with readers that could draw on to express himself or herself.

- (6) Another important reason that *make* Thai students are not ready to *learn* English online is they might not *understand* what they are *learning* because they do not *have* a right direction.

Table 5 The results of word concreteness analysis

z-score range	WC1 (n = 40)	WC2 (n = 40)
- 2.5 to < - 2.0	–	2
-2.0 to < - 1.5	2	6
- 1.5 to < - 1.0	6	19
-1.0 to < - 0.5	17	7
-0.5 to < 0.0	13	6
0.0 to < 0.5	2	–

Note: WC1 = word concreteness on writing task one, WC2 = word concreteness on writing task two

Action verbs (e.g. *make, learn, understand*) were found densely written in the passage (6) to describe the actions that the subject might do in each situation and expressed opinions with supported explanations of English learning experiences.

Referential cohesion

The results shown in Table 6 determine the degree to which words and ideas overlap across texts. Texts that contain a higher number of referential cohesion express distinct connections between ideas, so they are easier to read and comprehend.

Sixty percent of the texts from the writing task 1 (*n* = 24) were less than the mean value, indicating that the texts were high in referential cohesion that showed explicit connections between ideas in the texts. In contrast, a further analysis points to 45% of the texts from the writing task 2 (*n* = 18) contained a lower number of referential cohesion, indicating that the texts were relatively difficult to read and comprehend. Instead of using a variety of pronoun references, repeated pronouns were frequently found in the texts. The passage below illustrates this characteristic:

(7) For a decade that the internet has become accessible in Thailand. It makes *people's* life more convenient. For example, *people* do not have to go to the library when *they* need to find some information, or *they* can learn new lessons through online courses when *they* do not want to leave home.

The writer tends to reiterate what was mentioned in the previous sentences with the same words or pronouns. As shown in the passage, '*they*' was repeated three times. The last one of them in the extract "... when *they* do not want to leave home" shows the redundancy in the passage that seems to be a flaw because the writer failed to provide the noun *they* referred to as it was expected to appear.

Consider again by comparing the previous passage sample (7) with the following extract (8) from the writing task 2 written by the same student.

Table 6 The results of referential cohesion analysis

z-score range	RC1 (<i>n</i> = 40)	RC2 (<i>n</i> = 40)
- 1.5 to < - 1.0	3	-
- 1.0 to < - 0.5	10	3
-0.5 to < 0.0	11	15
0.0 to < 0.5	7	11
0.5 to < 1.0	7	3
1.0 to < 1.5	1	6
1.5 to < 2.0	-	1
2.0 to < 2.5	-	1
2.5 to < 3.0	-	-
3.0 to < 3.5	-	-
3.5 to < 4.0	1	-

Note: RC1 = referential cohesion on writing task one, RC2 = referential cohesion on writing task two

(8) As a result, *Thai students* do not gain adequate learning. For example, English is not our mother tongue language; therefore, *the students* cannot question about unknown knowledge when *they* do not comprehend *their* lesson.

Notice that in this passage (8), it extends the semantic domain of the concept of *Thai students* to include different lexical items like *the students* and *they* respectively. It can be assumed that the students showed an ability to adapt their writing to the requirements of the task at hand, thereby displaying an awareness of linguistic repertoires beyond that of the writing task 1. This is to suggest that uses of referential cohesion will reflect the degree to which words and ideas go beyond a text. Texts contain higher referential cohesion tend to show associations between ideas in the texts leading to more feasible to read.

Deep cohesion

The analysis of deep cohesion shows the degree to which the texts contain causal or intentional connectives in the sentences that help readers form a more coherent and deeper understanding of the causal events, processes, or actions in the texts. The results of deep cohesion analysis are presented in Table 7.

The results of the analysis show that a major range of the *z*-scores from the writing task 1 and 2 was 1.0–2.5, which is 62.5% (*n* = 25) and 72.5% (*n* = 29) respectively. The texts on the writing task 2 were at a higher and positive degree of *z*-scores than those of task 1, meaning that deep cohesion on the task 2 generates easier texts to comprehend. The following passage (9) on the second writing assignment illustrates the finding mentioned earlier.

(9) Firstly, education is essential as of the world today. In the present, the world is moving fast and keep on going. Technologies have developed *and* are more advanced as time passes. New inventions are being discovered rapidly. *Therefore*, it is important to keep the knowledge up to the standards of those new innovations.

The passage (9) shows that the writer attempted to use coordinating conjunction ‘*and*’ to connect the two different verbs in the present perfect and past participle forms as well as using conjunctive adverb ‘*therefore*’ to make a sequence of sentences. The

Table 7 The results of deep cohesion analysis

z-score range	DC1 (<i>n</i> = 40)	DC2 (<i>n</i> = 40)
0.0 to < 0.5	3	–
0.5 to < 1.0	6	2
1.0 to < 1.5	7	7
1.5 to < 2.0	11	11
2.0 to < 2.5	7	11
2.5 to < 3.0	2	4
3.0 to < 3.5	2	3
3.5 to < 4.0	2	2

Note: DC1 = deep cohesion on writing task one, DC2 = deep cohesion on writing task two

uses of conjunction create cohesion in the passage as well as extending and enhancing the text.

Besides, a majority of the texts from the first writing assignment expressed a fewer number of coordinating conjunctions (e.g. *and, but, so*) and conjunctive adverbs (e.g. *still, even though*) to supply cohesive ties across sentence boundaries. In addition, the z-scores reflected a fewer number of cohesive ties found in the overall texts causing relatively low in the associations of the lexical and grammatical structures, and the less sentence sequences to be understood as connected discourse (Halliday & Hasan, 1976).

In conclusion, the analyses provide evidence for the fact that the writers demonstrated linguistic flexibility across the texts that they produced. The participants wrote longer texts on the writing task 2 compared with those shorter texts on the task 1. The writers tend to develop more ideas around the writing topics as well as employing more sophisticated words in their texts on their second writing assignment. A majority of the participants could perform their abilities to use English vocabulary, extend their concepts, and express their complexity of English language. However, the findings also suggested that the participants may need to use a variety of lexical items to elaborate their writings, explicitly express ideas, and give new information to their readers.

In response to the research question 2, the results of the further analysis on correlations between the discourse components are presented in the following sections.

Multiple-variable correlations

Research question 2: Is there any significant correlation between those discourse components of the texts across the writing tasks?

The data from the 80 texts were examined to indicate the correlations between the results of analyses on the five discourse components.

As shown in Table 8, the referential cohesion was found associated with syntactic simplicity on the writing task 1 ($r = -.51, p = .00$), and on the writing task 2 ($r = -.39, p = .00$) respectively. One possible reason for the associations between the referential cohesion and complex syntactic sentences could be that more proficient writers tend to use deliberately syntactic structures by using cohesion to associate the lexicon with the

Table 8 Correlations between the components

Variable	1	2	3	4	5	6	7	8	9	10
1. Nar1	-									
2. Nar2	.227	-								
3. Syn1	-.262	-.111	-							
4. Syn2	-.156	.029	.614 ^a	-						
5. WC1	-.132	.145	-.144	.066	-					
6. WC2	-.004	-.469 ^a	.033	.087	.127	-				
7. RC1	.656 ^a	.274	-.511 ^a	-.167	.005	-.102	-			
8. RC2	.176	.482 ^a	-.411 ^a	-.397 ^a	.160	-.209	.444 ^a	-		
9. DC1	.099	-.020	.148	-.105	.114	-.023	-.159	-.006	-	
10. DC2	-.006	.472 ^a	.144	.091	.267	-.286	.032	.300	.076	-

Note. 1 = writing task 1, 2 = writing task 2, Nar = Narrativity, Syn = Syntactic Simplicity, WC = Word Concreteness, RC = Referential Cohesion, DC = Deep Cohesion

^aCorrelation is significant at the .01 level (2-tailed)

^bCorrelation is significant at the .05 level (2-tailed)

grammar in texts, generating sentence sequences as connected discourse of the texts (Ferretti & Lewis, 2019; Halliday & Hasan, 1976).

Amongst the variables on the second writing task, the narrativity was found significantly related to word concreteness ($r = -.46, p = .00$), referential cohesion ($r = .48, p = .00$), and deep cohesion ($r = .47, p = .00$). In particular, the participants' narrative flexibility was found associated with certain components that could influence the overall participants' writing performances of the argumentative texts they produced. Likewise, Allen et al. (2019) found that the narrative flexibility and referential cohesion were positively associated. Overall, linguistic flexibility and discourse components of texts were associated at a certain degree. The associations amongst the discourse components give another indication that the writers were able to tell their stories, expand, and connect their ideas through the concrete words and cohesion. More-skilled writers tend to have a greater number of working words to develop their writing, thereby expressing their ability to use sophisticated language and more diversity of words in their writings (Crossley, Kyle, & McNamara, 2016). The results of the analysis on syntactic simplicity from both writing tasks were found significantly related ($r = .61, p = .00$), indicating an interaction to a higher degree as syntactic sentences, and a shorter length of texts can predict the quality of the texts (Bulté & Housen, 2014; Polio & Shea, 2014).

Overall, the abilities to use linguistic features and discourse components of texts are varied by writers, meaning that to a certain degree their individual differences, prior knowledge, or writing ability can influence on writing performance. Besides, the results of the analyses also emphasize the importance of investigating a multi-dimensional perspective revolving around the linguistic features from basic levels to complicated discourse components. In doing this, writing teachers or researchers will have a better understanding of how to improve students' writing skill with more specific details in linguistic and syntactic dimensions.

Conclusions and future grounds for research

The findings of this study support the notion that the linguistic properties and the discourse components of texts are generally associated. The analyses focused on both natural language process to capture writing performances, and specifically computational linguistics ranging from basic text features to higher levels of measurement.

An examination of the discourse components of the overall texts shows relationships between the variables and yields important information for writing researchers and educators. In line with several studies, the discourse components cannot be isolated to indicate students' writing performance or proficiency. The connections between linguistic properties and writing process could be significant predictors of the students' writing performances.

The present study also shows that the implementation of automated tools in writing classes can be used to investigate students' texts in regards to the linguistic properties and the discourse components of the texts. The benefits of this use of an automated approach in this study were that the automated tool (Coh-Metrix) analyzed texts faster, more consistent, and on larger scales, overcoming constraints of a manual approach. In general, automated tools for writing analysis can repeatedly examine large textual data, thereby contributing validity and reliability. Besides, the automated tools yield great

potential in a substantial part of writing assessment as they can serve writing teachers as a second opinion for potentially accurate evaluation.

Future research may be integrated into rhetorical settings and the uses of various kinds of automated English writing tools to better evaluate and improve students' writing skill. In practice, writing teachers may use an automated tool in their classes to evaluate linguistic features of texts and demonstrate students' writing progress. Limitations of this study were the small number of participants. In order for researchers in the writing field to analyze larger samples and be able to generalize in ways that the researcher could not, better methods of recording revision changes may need to be explored.

Finally, technology has a great impact on teaching and learning L2 writing as it has been widely used to improve learners to higher levels of English writing skill. Technology may not fully replace the delivery of instruction and human interaction, as of now, it is potential shift from a content-based approach to increasing access to online educational resources and digital language platforms for teachers and students. Therefore, in practice, teachers should considerably select a teaching and learning approach and tools to improve students' writing skill and help them reach to their full potential in L2 writing.

Abbreviations

EFL: English as a Foreign Language; DC: Deep Cohesion;; L2: Second Language; NT: Narrativity; RC: Referential Cohesion; SS: Syntactic Simplicity; WC: Word Concreteness

Acknowledgements

I would like to gratefully and sincerely thank my colleagues from the Faculty of Liberal Arts at King Mongkut's Institute of Technology, Ladkrabang and the Faculty of Education at Ramkhamhaeng University for their insightful guidance and feedback. In particular, I would like to thank Dr. Sulawan Yottanoo for her expertise, assistance, and guidance in statistical interpretations and comments that greatly improved the manuscript. Finally, and most importantly, I would like to thank all research participants for giving me their valuable time and effort to contribute to the study.

Author's contributions

This study was conducted by a single author, and the only contributor is the author herself. The author read and approved the final manuscript.

Authors' information

Anongnad Petchprasert is an assistant professor in the Department of Curriculum and Instruction, Faculty of Education, Ramkhamhaeng University, Bangkok, Thailand. She received a Ph.D. in Curriculum and Instruction with a specialization in language education from Indiana State University, the United States. Her main research interests include corpora for linguistic use, natural language processing, second language acquisition, early childhood literacy, and TEFL/TESL.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Availability of data and materials

The data and materials will be available upon request.

Competing interests

The author declares that she has no competing interests.

Received: 15 April 2020 Accepted: 1 December 2020

Published online: 04 January 2021

References

- Allen, L. K., Jacovina, M. E., & McNamara, D. S. (2016). Computer-based writing instruction. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research*, (2nd ed., pp. 316–329). New York: Guilford Press.
- Allen, L. K., Likens, A. D., & McNamara, D. S. (2019). Writing flexibility in argumentative essays: a multidimensional analysis. *Reading and Writing*, 32, 1607–1634. <https://doi.org/10.1007/s11145-018-9921-y>.
- Aryadoust, V., & Liu, S. (2015). Predicting EFL writing ability from levels of mental representation measured by Coh-Metrix: a structural equation modeling study. *Assessing Writing*, 24, 35–58. <https://doi.org/10.1016/j.asw.2015.03.001>.

- Buckingham Shum, S., Sándor, Á., Goldsmith, R., Bass, R., & McWilliams, M. (2017). Towards reflective writing analytics: rationale, methodology and preliminary results. *Journal of Learning Analytics*, 4(1), 58–84. <https://doi.org/10.18608/jla.2017.41.5>.
- Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, 26, 42–65.
- Crossley, S. A., Bradfield, F., & Bustamante, A. (2019). Using human judgments to examine the validity of automated grammar, syntax, and mechanical errors in writing. *Journal of Writing Research*, 11(2), 251–270.
- Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The development and use of cohesive devices in L2 writing and their relations to judgments of essay quality. *Journal of Second Language Writing*, 32, 1–16.
- Crossley, S. A., Roscoe, R., & McNamara, D. S. (2014). What is successful writing? An investigation into the multiple ways writers can write successful essays. *Written Communication*, 31(2), 184–214.
- Deane, P. (2013). On the relation between automated essay scoring and modern views of the writing construct. *Assessing Writing*, 18, 7–24. <https://doi.org/10.1016/j.jasw.2012.10.002>.
- Elfenbein, A. (2011). Research in text and the uses of Coh-Metrix. *Educational Researcher*, 40(5), 246–248.
- Faigley, L., & Witte, S. (1981). Analyzing revision. *College Composition and Communication*, 32(4), 400–414.
- Ferretti, R. P., & Graham, S. (2019). Argumentative writing: theory, assessment, and instruction. *Reading and Writing*, 32, 1345–1357. <https://doi.org/10.1007/s11145-019-09950-x>.
- Ferretti, R. P., & Lewis, W. E. (2019). Best practices in teaching argumentative writing. In S. Graham, C. A. MacArthur, & J. Fitzgerald (Eds.), *Best practices in writing instruction*, (3rd ed., pp. 135–161). New York: Guilford Press.
- Graesser, A. C., McNamara, D. S., Cai, Z., Conley, M., Li, H., & Pennebaker, J. (2014). Coh-Metrix measures text characteristics at multiple levels of language and discourse. *Elementary School Journal*, 115, 211–229.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Hardy, J., & Friginal, E. (2016). Genre variation in student writing: a multi-dimensional analysis. *Journal of English for Academic Purposes*, 22, 119–131. <https://doi.org/10.1016/j.jeap.2016.03.0>.
- Haswell, R. H. (2000). Documenting improvement in college writing: a longitudinal approach. *Written Communication*, 17, 307–352.
- Humphreys, A., & Wang, R. J. H. (2017). Automated text analysis for consumer research. *Journal of Consumer Research*, 44(6), 1274–1306. <https://doi.org/10.1093/jcr/ucx104>.
- Kim, Y.-S. G., Gatlin, B., Otaiba, S. A., & Wanzek, J. (2018). Theorization and an empirical investigation of the component-based and developmental text writing fluency construct. *Journal of Learning Disabilities*, 51(4), 320–335. <https://doi.org/10.1177/0022219417712016>.
- Li, L. (2014). Contextual and cultural influence on the use of first person pronouns by Chinese learners of English. In D. D. Qian, & L. Li (Eds.), *Teaching and learning English in East Asian universities: global visions and local practices*, (pp. 302–322). Cambridge: Cambridge Scholars Publishing.
- MacArthur, C. A., Jennings, A., & Philippakos, Z. A. (2019). Which linguistic features predict quality of argumentative writing for college basic writers, and how do those features change with instruction? *Reading and Writing: An Interdisciplinary Journal*, 32, 1553–1574. <https://doi.org/10.1007/s11145-018-9853-6>.
- McNamara, D. S., Crossley, S. A., & Roscoe, R. (2013). Natural language processing in an intelligent writing strategy tutoring system. *Behavior Research Methods*, 45(2), 499–515.
- McNamara, D. S., Graesser, A. C., McCarthy, P., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge: Cambridge University Press.
- Polio, C., & Shea, M. C. (2014). An investigation into current measures of linguistic accuracy in second language writing research. *Journal of Second Language Writing*, 26, 10–27.
- Shermis, M. D., & Burstein, J. (2013). *Handbook of automated essay evaluation: current applications and future directions*. New York: Routledge.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1), 24–54. <https://doi.org/10.1177/0261927X09351676>.
- Ullmann, T. D. (2019). Automated analysis of reflection in writing: Validating machine learning approaches. *International Journal of Artificial Intelligence in Education*, 29, 217–257. <https://doi.org/10.1007/s40593-019-00174-2>.
- Varner, L. K., Roscoe, R. D., & McNamara, D. S. (2013). Evaluative misalignment of 10th-grade student and teacher criteria for essay quality: an automated textual analysis. *Journal of Writing Research*, 5, 35–59.
- Wilson, J., & Czik, A. (2016). Automated essay evaluation software in English language arts classrooms: effects on teacher feedback, student motivation, and writing quality. *Computers & Education*, 100, 94–109. <https://doi.org/10.1016/j.compedu.2016.05.004>.
- Zhu, W. (2001). Performing argumentative writing in English: difficulties, processes, and strategies. *TESL Canada Journal*, 19(1), 34–50.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.