Korean Journal of English Language and Linguistics, **Vol 24, July 2024, pp. 735-753**
DOI: 10.15738/kjell.24..202407.735

KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

ISSN: 1598-1398 / e-ISSN 2586-7474

http://journal.kasell.or.kr

Exploring the Impact of Cohesion on L2 Writing Proficiency: Beyond the Influence of the Lexical and Syntactic Complexity and the Use of N-Grams*

Xuemei Liang (Yanbian University, Hanyang University) · **Shinwoong Lee** (Hanyang University)

This is an open-access article distributed under the terms of the Creative Commons License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received: November 4, 2023
Revised: June 18, 2024
Accepted: July 16, 2024

Liang, Xuemei (First author)
Lecturer, Department of English
Yanbian University
977, Gongyuan Road, Yanji City, Jilin Province, PR China
Tel: 0086-433-273-2421
Email: liangxuemei@ybu.edu.cn

Lee, Shinwoong (Corresponding author)
Professor, Department of English Language and Literature
Hanyang University
222, Wangsimni-ro, Seongdong-gu, Seoul, 04763, Korea
Tel: 02-2220-0745
Email: shinwoonglee@hanyang.ac.kr

* This work was supported by the research fund of Hanyang University (HY-202100000003524).

ABSTRACT

Liang, Xuemei and Shinwoong Lee. 2024. Exploring the impact of cohesion on L2 writing proficiency: Beyond the influence of the lexical and syntactic complexity and the use of n-grams. *Korean Journal of English Language and Linguistics* 24, 735-753.

The current study aimed to investigate what cohesion indices could predict L2 writing proficiency and whether they could explain L2 writing proficiency significantly, even after controlling for the influence of the lexical and syntactic complexity and the use of n-grams. A Pearson correlation coefficient and a stepwise regression were conducted to examine the relationship between the use of cohesive devices and L2 writing proficiency by analyzing 1,200 essays written by Korean EFL university students. Four cohesion indices, including semantic relatedness and causal, temporal, and logical connectives, were found to be significant predictors of L2 writing proficiency. Furthermore, the results of hierarchical regression analysis revealed that even after controlling for the influence of lexical and syntactic complexity and the use of n-grams on the writing scores, the semantic relatedness was still significantly predictive of L2 writing proficiency, explaining 8% of the total variance of the writing scores. The findings suggested that semantic relatedness plays a crucial role in L2 writing proficiency, uniquely contributing to its development.

KEYWORDS

cohesion, semantic relatedness, lexical complexity, syntactic complexity, n-gram, L2 writing proficiency

© 2024 KASELL All rights reserved

1. Introduction

Over the past few decades, many L2 writing studies have investigated the relationship between the linguistic features of the texts written by L2 learners and their writing proficiency. These studies examined how language use can affect L2 writing proficiency, investigating such features as lexical and syntactic complexity (Chon et al. 2021, Kim and Crossley 2018, O'Leary and Steinkrauss 2022) and the use of multiword units including n-grams and lexical bundles (Garner et al. 2019, Stubbs 2007). Most of the studies have found that lexical and syntactic complexity can significantly contribute to L2 writing proficiency (Crossley 2020, Guo et al. 2013, Kyle and Crossley 2016) and revealed that high-scoring L2 essays have more high-frequency multiword units with higher association strength than low-scoring L2 essays (Garner et al. 2019, Zhang and Li 2021). In addition, the relationship between the use of cohesive devices and L2 writing proficiency was explored, and their use in L2 writing was found to be significantly associated with L2 writing proficiency in general (Crossley et al. 2016, Kim 2022, Lee 2021, Tywoniw and Crossley 2019, Yang and Sun 2012).

Meanwhile, several studies (Guo et al. 2013, Kim and Crossley 2018, Zhang et al. 2022) attempted to examine the influence of lexical, syntactic, and cohesive measures on L2 writing proficiency together. It was largely found that lexical and syntactic features and some cohesive measures could be significantly predictive of L2 writing proficiency. However, although the findings of those studies (Crossley and McNamara 2012a, Guo et al. 2013, Ryu 2020, Zhang et al. 2022) mostly confirmed the contribution of the use of cohesive devices to L2 writings along with lexical and syntactic complexity of L2 writings, it appears that some of the measures of lexical and syntactic complexity utilized in previous studies (Crossley and McNamara 2012a, Crossley et al. 2016, Guo et al. 2013, Yang and Sun 2012) may have overlapped with those of cohesion showing interdependency among the measures. Thus, it is possible that the findings of those studies may not have adequately captured the unique contribution of cohesion to L2 writing proficiency, as its influence on L2 writings is possibly intertwined with that of lexical and syntactic complexity. In this vein, it may be necessary to reexamine how much cohesive devices in L2 writing can contribute to L2 writing proficiency when controlling for the influence of lexical and syntactic complexity. Indeed, few studies have investigated the contribution of cohesion to L2 writing proficiency after controlling for the impact of other linguistic variables. Thus, the primary purpose of the current study was to examine what cohesive devices could predict L2 writing proficiency and whether there existed a unique contribution to L2 writing proficiency after controlling for the influence of lexical and syntactic complexity on L2 writing.

2. Literature Review

2.1 Linguistic Complexity and L2 Writing Proficiency

Previous studies on the relationship between linguistic complexity and L2 writing proficiency have often investigated how lexical and syntactic complexity relate to L2 writing proficiency (e.g., Chon et al. 2021, Kim and Crossley 2018, O'Leary and Steinkrauss 2022, Yoon 2021). Lexical complexity refers to the level of difficulty or sophistication of the vocabulary used in a text, often involving an analysis of the complexity of vocabulary and word choices in writing. It has been mostly measured through three lexical aspects: lexical density, lexical sophistication, and lexical variation (Lu 2012, Read 2000) and is primarily used for linguistic research and communication analysis (Kyle et al. 2018, Lu 2012, Read 2000).

© 2024 KASELL All rights reserved

Lexical density refers to the ratio of lexical words (e.g., nouns, verbs, adjectives, and adverbs) to all words in a text (Lu 2012, Nasseri and Thompson 2021). A higher lexical density indicates a greater concentration of content words, leading to more information-rich writing (Biber and Gray 2016), and it has been generally found that lexical density is significantly associated with L2 writing proficiency (Kim 2014, Nasseri and Thompson 2021). Moreover, lexical sophistication refers to "the proportion of relatively unusual or advanced words in the learner's text" (Read 2000, p. 203). It has been revealed that lexical sophistication could significantly affect writing proficiency in L1 (Douglas 2013) and L2 (Crossley and McNamara 2012a). Crossley and McNamara (2012a) reported that lexical sophistication accounted for 29% of the variance of L2 writing proficiency. Additionally, lexical variation or diversity refers to the degree of vocabulary diversity exhibited in language production, typically measured by a type-token ratio (TTR) and its variants. Previous studies on lexical variation and L2 writing proficiency (Crossley and McNamara 2012a, Kim 2014) suggested that it significantly contributes to L2 writing proficiency.

Also, syntactic complexity refers to the sophistication of sentential and clausal structures in a text (Lu 2010, 2011). It has been argued that proficient writers tend to produce more complex sentences (Lu 2010, 2011, McNamara et al. 2010). Lu (2010) examined fourteen syntactic complexity measures with Chinese EFL learners' writings, and the findings indicated that more proficient L2 writers tended to produce more complex nominals, coordinated phrases, and longer clauses. Along the same line, Ai and Lu (2013) compared the essays written by native speakers of English and Chinese EFL learners by measuring syntactic complexity (i.e., length of production, subordination, coordination, and phrasal complexity). They found that there existed significant differences between the native speakers and the EFL learners in all four categories, while the high-proficiency EFL learners were significantly higher than the low-proficiency EFL learners in those measures. Furthermore, Lee et al. (2021) examined the measures of lexical and syntactic complexity in advice reports and letters written by Chinese learners of English, and they found that such syntactic measures as the mean length of clauses, coordinate phrases per clause, and complex nominals per clause were significantly predictive of L2 writing proficiency.

In the meantime, multiword units such as n-grams (Garner et al. 2019, Stubbs 2007) and lexical bundles (Bestgen 2017, Bestgen and Granger 2014) have been examined to see if they can contribute to L2 writing proficiency. N-grams refer to a sequence of continuous words with an n-number of words from a spoken or written text, and lexical bundles are the sequences of words that occur together more often than would be expected by chance (Bestgen 2017). With the importance of multiword units in language production widely recognized in both spoken and written language, multiword units such as n-grams and lexical bundles have become an important language unit in the investigation of language production (Ädel and Erman 2012, Cortes 2013, Garner et al. 2019, Salazar 2014). Accordingly, several studies attempted to investigate the relationship between using multiword units and L2 writing proficiency.

Garner et al. (2019) investigated the influence of n-grams on L2 writing proficiency, and they found that the use of n-grams, especially the proportion of n-grams, could be a significant predictor of human-rated writing scores. Similarly, Zhang and Li (2021) investigated the impact of n-gram measures (range, frequency, and association strength) on the quality of L2 writing. The results revealed that various n-gram measures, including spoken bigram range, spoken bigram proportion 10K, academic trigram frequency, and spoken bi-and-trigram association strength (i.e., Delta P and MI), were significantly predictive of L2 writing proficiency. Also, Kim and Kessler (2022) investigated the use of lexical bundles in 120 academic essays of Chinse EFL learners by examining the relationship between the use of lexical bundles and the overall quality of the writings. The findings indicated quantitative and qualitative differences in using 3-word bundles between high- and low-level essays.

In general, it was found that lexical and syntactic complexity play a crucial role in shaping the quality of L2 writing, and a varied lexicon and complex syntax often signify advanced writing skills, allowing writers to express

© 2024 KASELL All rights reserved

ideas effectively. Also, the findings of the studies on the relationship between the use of multiword units and L2 writing suggested that the knowledge of multiword units such as n-grams and lexical bundles is significantly predictive of L2 writing proficiency. To sum up, it was suggested that lexical and syntactic complexity and the knowledge of multiword units significantly contribute to L2 writing proficiency.

2.2 Cohesion and L2 Writing Proficiency

Cohesion is defined as "the presence or absence of linguistic cues in the text that allows the reader to make connections between the ideas in the text" (Crossley et al. 2016, p. 2), and it has been reported that not only L1 writers but also L2 writers utilize lexical, syntactic, and textual devices to have connectedness and organization throughout a text (Tywoniw and Crossley 2019). In this vein, cohesion has been recognized as one of the dimensions of L2 writing to be assessed (Crossley and McNamara 2012a, Crossley et al. 2016, Guo et al. 2013, Kim and Crossley 2018, Yang and Sun 2012). Cohesion, in general, is categorized into three types: local, global, and text cohesion (Crossley et al. 2016). Local cohesion refers to the use of cohesive devices at a sentence level, such as the use of connectives (e.g., *because, so*) and lexical overlaps between sentences. Global cohesion denotes the use of cohesive devices for lexical and semantic overlaps at a paragraph level, and text cohesion indicates the occurrence of cohesion at a whole text level.

As the influence of cohesion on the quality of writing has been recognized, several studies attempted to examine the influence of the use of cohesive devices on L2 writing proficiency. These studies found that the influence of cohesion on L2 writing quality exhibited variation depending on types of genres (Ryu 2020, Zhang et al. 2022), task types (Guo et al. 2013, Kim and Crossley 2018, Tywoniw and Crossley 2019), and L1 backgrounds (Crossley and McNamara 2012b). For instance, Ryu (2020) investigated predictive linguistic features in L2 writing, focusing on the genres of argumentation and narration. She found that the use of cohesive devices was more indicative of L2 writing proficiency in the analysis of argumentative essays than in the analysis of narrative essays. Also, Tywoniw and Crossley (2019) examined the impact of cohesion on L2 writing proficiency by comparing two types of writing tasks: integrated writing and independent writing. The result of regression analysis showed that the lexical overlap of function words had a positive impact on both types of writing tasks. Meanwhile, Crossley and McNamara (2012b) reported the differences among L2 learners of Czech, Finnish, German, and Spanish in the use of cohesive devices. Spanish learners were found to use causal cohesions and noun overlaps most often, while Czech learners utilized given information and temporal cohesion in their essays.

Meanwhile, the findings of some of the studies on the relationship between the use of cohesive devices and L2 writing proficiency were somewhat inconclusive. Guo et al. (2013) investigated the impact of lexical, syntactic, and cohesive elements on L2 writing proficiency in two different writing tasks (i.e., integrated and independent tasks). Their study revealed a multifaceted relationship between the use of cohesive devices and writing proficiency. While a positive relationship existed between cohesion and proficiency in integrated tasks, a negative relationship was observed in independent tasks. This underscored that task type could play a significant role in determining the relationship between cohesion and L2 writing proficiency. In addition, Crossley and McNamara (2012a) probed the effect of linguistic sophistication and cohesion on L2 writing proficiency, and the result of the study did not show a significantly positive relationship between L2 writing proficiency and the use of cohesive devices.

It should also be noted that in some of the previous studies (Crossley and McNamara 2012a, Crossley et al. 2016, Guo et al. 2013, Yang and Sun 2012) that explored the impact of cohesion on L2 writing, the measurement of cohesion may partially overlap with the evaluation of lexical and syntactic complexity. As previously mentioned, increased lexical diversity, one of the measures of lexical complexity, can lead to reduced lexical overlap, which

© 2024 KASELL All rights reserved

is one of the measures of cohesion, and this implies that there may exist a negative relationship between lexical complexity and cohesion. In contrast to a potential inverse relationship between lexical complexity and cohesion, syntactic complexity and cohesion appear to have a positive relationship. For instance, the sentence, "These reasons may not be very convincing, but I firmly believe that spiritual satisfaction is really a happy thing" (Yang and Sun 2012, p. 39), is syntactically more complex than two separate sentences without "but" in terms of production length (e.g., number of words per clause, sentence, and T-unit) and sentence complexity (e.g., number of clauses per sentence). As such, syntactic complexity might be related to the use of connectives (e.g., but, and, thus, etc.), which is one of the measures of cohesion of the text. Furthermore, such n-grams as "first of all," "as a result," and "for example" that function as connectives can influence the cohesion of a text as well. In this vein, it may be possible to posit that the assessment of cohesion may intersect with the measurements of lexical complexity, syntactic complexity, and n-grams to some extent.

2.3 The Current Study

As aforementioned, there have been several studies that examined the relationship between the use of cohesive devices and the quality of L2 writing, but few studies have looked into whether there exited a unique contribution of cohesion to L2 writing proficiency even after controlling for the influence of lexico-grammatical complexity and the use n-grams. Given that, the primary purpose of the current study was to identify the variables of cohesion that are significantly predictive of L2 writing proficiency and investigate if there exists a unique contribution of those variables to L2 writing proficiency even when controlling for the influence of lexical and syntactic complexity and the use of n-grams. The research questions are as follows:

1) What cohesion indices are significantly associated with and predictive of L2 writing proficiency?
2) How much can the use of cohesive devices contribute to L2 writing proficiency after controlling for the influence of the lexical and syntactic complexity and the use of n-grams?

3. Method

3.1 Data Collection

The data used for the analysis were the essays written by the students who were newly admitted into the university located in Seoul, Korea, in 2020. The writing test, which was one of the sub-components of the General Test of English Language Proficiency (G-TELP), was administered to freshmen right before the spring semester began in 2020, and it served as a placement test for academic English courses at the university. The International Testing Services Center developed G-TELP as a tool for assessing the English proficiency of non-native speakers of English, and it is designed to evaluate English proficiency in general, including listening, reading, speaking, and writing skills, and has been used in many educational institutions, companies, and organizations for evaluating English proficiency with diverse purposes (Kelly 1988, Kim and Seol 2021, Utsunomiya et al. 2016). In particular, the G-TELP writing component is designed to evaluate an individual's ability to communicate in written form, with a focus on vocabulary, grammar, sentence structure, coherence, and organization of writing. The examination, conducted through a computer interface, encompassed sections on listening, reading, and writing, all of which were to be completed within 150 minutes. Only the writing segment derived from the G-TELP was utilized in the

© 2024 KASELL All rights reserved

current study. The participants in the study were required to choose one out of four topics and write an essay on the selected theme within 30 minutes. The essay prompts, labeled A to D, are detailed in Table 1. The use of any reference materials by the students was not allowed during the test.

Table 1. Topics of the Essays

	Prompts
Topic A	Should lawmakers restrict gun owner rights? Please write your opinion on gun control with specific reasons.
Topic B	Do you agree or disagree with the idea of taking a gap year between high school and university? Why?
Topic C	Do you think professional sports players' salaries are too high?
Topic D	Do you agree with the idea that using cell phones from a young age decreases the desire to go outside and have face-to-face interaction?

The writings of the students were assessed by the official raters from the G-TELP. According to the assessment guidelines of writing, expert raters assess the written texts based on various aspects, including style, grammar, vocabulary, organization, and substance. Then, the essays are categorized into 11 proficiency levels based on the average ratings of those five areas (refer to Table 2 and http://www.itsc-group.com/writing6.php for detailed rubrics of evaluation and descriptions).

Table 2. Description of the Proficiency Levels in G-TELP Writing

Level	Score	Description
1	95~100	Writers are able to express themselves with ease in all situations, whether familiar or unfamiliar.
2	85~94	Writers are able to communicate their ideas effectively in nearly all situations.
3	75~84	Writers are usually able to communicate their ideas effectively in nearly all situations.
4	65~74	Writers are generally able to communicate their ideas in most situations.
5	55~64	Writers are generally able to communicate their ideas in common situations but may occasionally have trouble when dealing with unfamiliar or uncommon events.
6	45~54	Writers are generally able to communicate their ideas in common situations but are sometimes unable to effectively express themselves when dealing with unfamiliar or uncommon situations.
7	35~44	Writers generally have a difficult time communicating their ideas in common situations and are often unable to effectively express themselves when dealing with unfamiliar or uncommon situations.
8	25~34	Writers usually have a difficult time expressing their ideas in common situations and are frequently unable to effectively respond when dealing with unfamiliar or uncommon situations.
9	15~24	Writers frequently have a difficult time communicating their ideas in common situations and are almost always unable to effectively respond when dealing with unfamiliar or uncommon situations.
10	5~14	Writers almost always have a difficult time expressing their ideas, even in common or familiar situations.
11	0~4	Writers may exhibit a vocabulary of a handful of memorized words and isolated phrases only. They are unable to express themselves in a meaningful way.

© 2024 KASELL All rights reserved

All the texts of the students' writings, as well as the student information including writing topics and their evaluation scores, were saved for analysis. The essays with a word count below the minimum requirement of 100 words were excluded from the analysis to avoid any potential problems that may arise when using automated text analysis tools with short texts, such as skewed results (Crossley and McNamara 2013, Garner et al. 2019). Also, the essays classified as level 1 and level 10 were excluded from the analysis as they were exceedingly uncommon (Garner et al. 2019). The number of essays that remained after the exclusions was 662 for Topic A, 886 for Topic B, 366 for Topic C, and 436 for Topic D, making a total of 2350 essays. To ensure that each topic was equally represented, 300 essays were chosen randomly from each topic, resulting in a total of 1,200 essays that were included for data analysis in the current study (See Table 3). The analysis of variance (ANOVA) revealed that there was no significant difference in the mean scores of the writing assessment across the four topics (p =.10). Table 4 shows descriptive statistics of the number of words from level 2 to 9.

Table 3. Number of Essays for Each Topic by Proficiency Level

	Level								
	2	3	4	5	6	7	8	9	Total
Topic A	9	34	42	67	55	60	19	14	300
Topic B	3	31	56	88	66	49	4	3	300
Topic C	8	25	49	92	63	46	14	3	300
Topic D	7	38	39	82	63	46	18	7	300

Table 4. Descriptive Statistics of Word Counts by Proficiency Level

	Level							
N of Word	2	3	4	5	6	7	8	9
Median	327	309	310	263	206	191	177	143
Mean	375	318	293	265	228	212	200	180
SD	118	96	71	72	87	85	79	115
Min.	177	117	106	100	101	101	100	101
Max.	634	975	550	555	913	627	396	645
Sum	10125	40738	54459	87135	56322	42534	11019	4853

3.2 Tools for Analysis and Indices of Measurement

To measure the cohesive features of the essays, the current study adopted cohesion indices from Coh-Metrix (Crossley and McNamara 2012a, Graesser et al. 2004,), which has been widely used to measure cohesion and linguistic sophistication of language productions (Chon et al. 2021, Crossley and McNamara 2012a). Among more than one hundred indices provided by Coh-Metrix, the current study chose 20 indices pertaining to local, global, and text cohesions, covering lexical overlap, semantic similarity, connectives, causal and temporal cohesion that have been used widely in the previous studies of cohesion in L2 writing (Crossley and McNamara 2012a, Crossley et al. 2016).

As shown in Table 5, five indices of lexical overlaps were utilized in the current study, including noun overlap, argument overlap, stem overlap, content word overlap, and anaphor overlap at a local (between adjacent sentences) and global level (within a paragraph). Noun overlap quantifies the frequency of shared noun forms, and argument

© 2024 KASELL All rights reserved

overlap measures how often two sentences share nouns with common stems. Meanwhile, stem overlap assesses the frequency of nouns that share a common semantic morpheme, and content word overlap includes the overlaps of nouns, verbs, adjectives, and adverbs.

In addition, the two types of semantic similarity, including semantic overlap and LSA given/new were measured utilizing Latent Semantic Analysis (LSA), one of the functions built in Coh-Metrix. Coh-Metrix provides the measures of semantic overlap between adjacent sentences (local), all sentences in a paragraph (global), and adjacent paragraphs (text). Meanwhile, LSA given/new pertains to the proportion of given information compared to new information in a text, and less-given information indicates lower cohesion (McNamara et al. 2014). Regarding connectives, the frequencies of causal connectives (e.g., *because, so*), contrastive connectives (e.g., *although, whereas*), additive connectives (e.g., *and, moreover*), temporal connectives (e.g., *first, until*), and logical connectives (e.g., *or, if, then*) used in the texts were measured. Also, temporality was measured by the repetition of aspect and tense, as well as their combination.

Table 5. Cohesion Indices Selected from Coh-Metrix (Graesser et al., 2004)

Cohesion type	Cohesion measures	Cohesion levels	N
Lexical overlap	noun overlap	local, global	2
	argument overlap	local, global	2
	stem overlap	local, global	2
	content word overlap	local, global	2
	anaphor overlap	local, global	2
Semantic similarity	semantic overlap	local, global, text	3
	LSA given/ new	global	1
Connectives	causal connectives	local	1
	contrastive connectives	local	1
	temporal connectives	local	1
	additive connectives	local	1
	logical connectives	local	1
Temporality	tense and aspect repetition	text	1

In measuring the lexical complexity of the writings, Lexical Complexity Analyzer (LCA; Ai and Lu 2010, Lu 2012) was used in the current study. It is a computational tool that calculates 25 measures of lexical complexity, encompassing the levels of lexical density (1 index), lexical sophistication (5 indices), and lexical variation (19 indices). Lexical density refers to the proportion of content words, while lexical sophistication pertains to the proportion of sophisticated words in texts. In addition, lexical variation measures numerous variants of the Types and Type-Token Ratio (TTR) of content words. Meanwhile, the syntactic complexity of the essays was measured by L2 Syntactic Complexity Analyzer (L2SCA; Lu 2010), which measures 14 indices of syntactic complexity of phrases, clauses, and sentences. The current study employed all of the 14 indices provided by L2SCA, covering five categories: length of production, sentence complexity, subordination, coordination, and phrasal complexity (See Appendix 1 for more information about the indices of lexical and syntactic complexity).

In addition, Tool for the Automatic Analysis of Lexical Sophistication 2.2 (TAALES 2.2; Kyle and Crossley 2015) was utilized to analyze n-grams in the current study. It provides frequency, range, proportion, and association strength of bigrams and trigrams based on the data from the Corpus of Contemporary American English (COCA;

© 2024 KASELL All rights reserved

Davies 2009). The current study utilized 86 indices of n-grams from academic and spoken subsections of COCA, which were built in TAALES (See Appendix 2 for more information).

3.3 Statistical Analysis

First, 20 cohesion indices were checked for normality, and a Pearson correlation coefficient was calculated between the scores of each cohesion index and the writing scores. Then, the cohesion indices that had a significant correlation coefficient over 0.1[1] (i.e., a small effect size, Cohen 1988) with the writing scores and that were not multicollinear with each other were selected[2]. Next, a linear stepwise regression was conducted with the writing scores as a dependent variable and the selected cohesion indices as independent variables. It was attempted to find what variables of cohesion were significantly predictive of the writing scores.

Then, in order to investigate the unique contribution of cohesion to L2 writing proficiency after controlling for the influence of lexical and syntactic complexity and the use of n-grams, a hierarchical regression[3] was employed. In step 1, the indices of lexical and syntactic complexity and n-grams that were found to be significantly associated with the L2 writing scores were entered as control variables. In selecting control variables, the correlations between the indices of linguistic complexity (i.e., the indices of lexical and syntactic complexity and n-grams) and writing scores were obtained, and only the variables whose correlation with the writing scores were significant and over r =.10 (Cohen 1988) and showing no multicollinearity between them were selected. In step 2, the variables of cohesion that were significantly predictive of L2 writing proficiency were entered to obtain the unique contribution of cohesion to L2 writing proficiency when controlling for the influence of lexical and syntactic complexity and the use of n-grams.

4. Results

4.1 What Indices of Cohesion Are Associated with and Predictive of L2 Writing Proficiency Significantly?

The results of Pearson's correlation coefficient showed that nine out of 20 cohesion devices[4] were significantly associated with the writing scores, among which causal, logical, adversative, and negative connectives and temporality were negatively associated with the writing scores, while semantic similarity (global), LSA given/new (global), stem overlap (local) and temporal connectives were positively correlated with the writing scores. Among the indices that were found to be significantly associated with the writing scores, only LSA given/new (global),

[1] Garner et al. (2019), Zhang and Li (2021) adopted this criterion (r > .01) in their studies of the relationship between linguistic complexity and L2 writing.

[2] If two indices exhibited a strong correlation with each other, with a correlation coefficient r > .70, only the index that displayed a higher correlation with the writing score was kept.

[3] According to Eddington (2015, p. 104), "hierarchical regression is used to control for the influence of one or more variables while measuring how much variance one or more additional variables account for." He claimed that the independent variables in the regression model might overlap in the variations, and "variables that are significant in one model can become insignificant (or vice versa) when another variable is included."

[4] Descriptive statistics of 20 cohesion indices in the current study, including mean, standard deviation, minimum, and maximum values of each index, is provided in Appendix 3.

© 2024 KASELL All rights reserved

temporal connectives, causal connectives, and logical connectives were found to have a significant correlation with the writing scores ($r > .10$) and didn't show multicollinearity (See Table 6).

Table 6. Cohesion Indices That Are Significantly Correlated with Writing Scores

Variables	*r*	*p*
LSA given/new (global)	.16	< .01
Temporal connectives	.11	< .01
Stem overlap (local)	.09	< .01
Semantic overlap (global)	.08	< .01
Causal connectives	-.13	< .01
Logical connective	-.12	< .01
Adversative connectives	-.06	< .05
Negative connectives	-.09	< .01
Temporality	-.09	< .01

Table 7 below displays the results of linear stepwise regression with the four indices of cohesion (i.e., LSA given/new, temporal connectives, causal connectives, and logical connectives) that were found to be significantly associated with the writing scores with little multicollinearity. The results showed that all of those four indices of cohesion were found to be significant predictors of the writing scores, resulting in a significant model, F (4, 1195) = 18.99, $p < .01$, $R^2 = 0.06$. The regression model accounted for 6% of the total variance of the writing scores, and LSA given/new was found to be the strongest predictor, followed by causal connectives, temporal connectives, and logical connectives. The variables such as LSA given/new and temporal connectives were positively predictive of L2 writing scores, while such indices as causal and logical connectives were negatively predictive of L2 writing scores. Meanwhile, all of the tolerance scores were under 1.00, and the variance inflation factors (VIF) were under 1.50, showing little multicollinearity among the variables.

Table 7. The Results of Stepwise Regression

Predictors	*B*	*t*	R^2	*Adjusted* R^2	*F*	R^2 *change*
LSA given/new	49.15	5.97**	0.024	0.023	29.80**	0.024
Causal connectives	-0.08	-2.31*	0.042	0.041	26.61**	0.018
Temporal connectives	0.18	4.18**	0.054	0.052	22.86**	0.012
Logical connectives	-0.08	-2.65**	0.06	0.058	18.99**	0.006

* $p < .05$, ** $p < .01$

4.2 The Unique Contribution of Cohesion to L2 Writing Proficiency

In order to identify the unique contribution of cohesion to L2 writing proficiency, it was attempted to select the control variables of lexical and syntactic complexity and n-grams. All the control variables underwent Pearson correlation and multicollinearity checks. Only the variables that were significantly correlated with writing scores ($r > .10$) and showed no multicollinearity between each other ($r < .70$) were chosen for hierarchical regression analysis. Among 25 indices of lexical complexity analyzed by Lexical Complexity Analyzer (LCA), corrected

© 2024 KASELL All rights reserved

TTR (r = .34) and corrected verb sophistication I (r = .14) were found to be significantly correlated with the writing scores without multicollinearity. Meanwhile, syntactic complexity was measured by L2 Syntactic Complexity Analyzer (L2SCA), and it was revealed that two indices of syntactic complexity, including coordinate phrase per clause (r = .12) and complex nominal per clause (r = .14) out of 14 indices were significantly associated with the writing scores without multicollinearity (see Appendix 1 for detailed information). Regarding n-grams, by employing Tool for the Automatic Analysis of Lexical Sophistication (TAALES), three indices of n-grams, including academic trigram proportion (r = .23), spoken trigram 2 ΔP (r = .20), and spoken bigram MI (r = .17), were found to have a significant correlation with the writing scores among the 43 indices (See Appendix 2 for detailed information). Table 8 shows Pearson's correlation coefficients between the selected indices of lexical complexity, syntactic complexity, and n-grams and the writing scores. All the indices of lexical and syntactic complexity and n-grams selected showed a normal distribution and little multicollinearity among the variables.

Table 8. Indices That Are Significantly Correlated with Writing Scores

Variables	Type	*r*	*p*
Corrected TTR	Lexical complexity	.34	< .01
Corrected verb sophistication I	Lexical complexity	.14	< .01
Coordinate phrases per clause	Syntactic complexity	.12	< .01
Complex nominals per clause	Syntactic complexity	.14	< .01
Academic trigram proportion	N-grams	.23	< .01
Spoken trigram 2 ΔP	N-grams	.20	< .01
Spoken bigram MI	N-grams	.17	< .01

Once the control variables of syntactic and lexical complexity and n-grams were chosen, a hierarchical multiple regression was conducted to see if a unique contribution of cohesion to L2 writing proficiency existed. In step 1, the indices of lexical complexity (i.e., corrected TTR and corrected verb sophistication I), syntactic complexity (i.e., coordinate phrases per clause and complex nominals per clause), and n-grams (i.e., academic trigram proportion, spoken trigram 2 Delta P, and spoken bigram MI) that were found significantly associated with the L2 writing scores were entered as control variables. Among the entered variables, corrected TTR (t = 9.23), academic trigram proportion (t = 6.13), spoken trigram 2 ΔP (t = 2.42), spoken bigram MI (t = 3.42), and coordinate phrases per clause (t = 2.29) were found to be significant predictors of L2 writing scores, and they could form a significant model to predict L2 writing proficiency explaining about 17% of the total variance of the writing scores, F (7, 1192) = 35.45, p < .01, R^2 = 0.172 (see Table 9).

In step 2, the four variables of cohesion that were found to be significantly predictive of the writing scores (i.e., LSA given/new, causal connectives, temporal connectives, and logical connectives) were entered, and they were found to explain an additional 8% of the total variance of L2 writing, F (11, 1188) = 36.12, p < .01, R^2 = 0.251 (See Table 9). However, only LSA given/new was found to be a significant predictor of the writing score. All of the tolerance scores were lower than 1, and variance inflation factor (VIF) scores were under 2.

© 2024 KASELL All rights reserved

Table 9. Results of Hierarchical Regression Analysis

Models	Predictors	*B*	*t*	R^2	*Adjusted* R^2	*F*	R^2 *change*
Step 1	Corrected TTR	0.27	9.23**	0.172	0.167	35.45**	0.172
	Corrected verb sophistication I	0.02	0.60				
	Academic trigram proportion	0.17	6.13**				
	Spoken trigram 2 Delta P	0.07	2.42*				
	Spoken bigram MI	0.10	3.42**				
	Coordinate phrases per clause	0.07	2.29*				
	Complex nominals per clause	0.01	0.19				
Step 2	LSA given/new	0.29	10.94**	0.251	0.247	36.12**	0.08
	Causal connectives	0.01	0.14				
	Temporal connectives	0.04	1.44				
	Logical connectives	-0.05	-1.58				

* $p < .05$, ** $p < .01$

5. Discussion

The current study aimed to identify the cohesion devices that can explain L2 writing proficiency significantly and to see whether they can contribute to L2 writing proficiency significantly even after controlling for the influence of the lexical and syntactic complexity and the use of n-grams. The results of stepwise regression analysis revealed that four cohesion indices (i.e., LSA given/new, causal connectives, temporal connectives, and logical connectives) could predict L2 writing proficiency significantly. The strongest predictor of L2 writing proficiency was LSA given/new, and it alone explained 2.4% of the total variance of L2 writing proficiency, which indicates that high-quality essays utilized more given information compared to new information. The findings of the current study are compatible with those of previous studies (Crossley and McNamara 2009, Guo et al. 2013), which reported that more proficient English writers tend to use words and phrases that are semantically relevant to the previous ones.

Another significant predictor of L2 writing proficiency was the use of causal connectives, which was negatively associated with the writing scores, accounting for 1.8% of the variance of the essay scores. The results indicated that more proficient L2 learners tended to use fewer causal connectives in their writing. The findings are in line with those of previous studies (Crossley et al. 2014, Jung et al. 2019), which showed that the overuse of causal connectives was negatively associated with L2 writing proficiency. This may be because less skilled writers are overly reliant on connectives to make logical connections between ideas clear, while more skilled writers can achieve their effectiveness through means other than overt causal connectives, possibly utilizing more varied vocabulary and structures to convey relationships between ideas.

Furthermore, temporal connectives were found to be a significant predictor, accounting for 1.2% of the total variance of the writing scores. It indicated that L2 learners with a higher essay score used more temporal connectives in their writing. The findings of the current study were different from those of previous studies, most of which showed no significant relationship between them (Crossley and McNamara 2012a, McNamara et al. 2010). This may be because the use of temporal connectives (e.g., *firstly, secondly, lastly*, etc.) could be perceived as an effective way to create a logical and chronological flow for Korean EFL learners, especially when they are asked to write a formal essay within a limited time. This might also be related to the L2 writing instructions they

© 2024 KASELL All rights reserved

received, where the use of temporal connectives was explicitly taught and possibly encouraged to establish a logical flow of L2 writings.

In addition, the use of logical connectives was found to be a significant predictor of L2 writing proficiency, explaining 0.6% of the total variance of the essay ratings. The negative relationship between the frequency of logical connectives and L2 writing scores indicated that high-scoring essays were likely to have fewer logical connectives compared to low-scoring essays. The findings were similar to those of the previous studies, which reported that high-scoring essays typically included fewer logical connectives in L2 writings (Crossley and McNamara 2012a, Tywoniw and Crossley 2019, Zhang et al. 2022).

Regarding the results of the hierarchical regression analysis, it was found that four cohesion indices accounted for 8% of the total variance of L2 writing proficiency. However, LSA given/new was the only significant predictor of L2 writing proficiency. It is notable that there was a rise in the variance of L2 writings explained by cohesion, increasing from 6% in the stepwise regression model to 8% in the hierarchical regression model when controlling for the influence of lexical and syntactic complexity and use of n-grams. This result may support the claim that the measurement of cohesion is partially affected by the assessment of lexical and syntactic complexity. In other words, the contribution of cohesive devices to L2 writings may be somewhat obscured by other linguistic features.

In general, the findings of the present study about the relationship between the use of cohesive devices and L2 writing proficiency were in line with those of previous studies (Crossley et al. 2014, Jung et al. 2019, Tywoniw and Crossley 2019, Zhang et al. 2022) supporting the claim that there exists a significant relationship between the use of cohesive devices and L2 writing proficiency. However, it should be noted that the variance of the writing scores that could be accounted for by the use of cohesive devices in the current study was low in comparison with the findings of previous studies. This may be because the length of the essays was too short to capture the variance of the frequency of cohesive and other linguistic devices adequately, which was mainly due to the fact that the participants were given only 30 minutes for essay writing. Therefore, in order to have more valid and reliable statistical results, a larger number of words in the essays should be guaranteed by providing L2 writers with sufficient time for essay writing. It should also be noted that the four cohesion indices, all of which were found to be significant predictors in the stepwise regression, ended up with only one index (i.e., LSA given/new) in the hierarchical regression. This indicates that the high-scoring L2 essays in the current study largely utilized the given information in their essay writing, which is compatible with the findings of previous studies (Crossley and McNamara 2009, Guo et al. 2013), whereas the other cohesive devices, including causal, temporal, and logical connectives appeared to be partialled out by other variables and have less influence on the quality of L2 writing.

The findings of the current study could provide implications for automated essay scoring (AES) systems such as e-rater[5] that have been used for the writing assessment of standardized English tests such as TOEFL and GRE. The advantages of the AES systems include efficient and consistent grading and the ability to provide immediate feedback to test-takers. However, most of them have very limited indices of cohesion in assessing L2 writing proficiency yet, even if they have been found to be a significant predictor of L2 writing proficiency persistently. Thus, the inclusion of adequate indices of cohesion in the AES system would contribute to the enhancement of the validity of the automated writing system.

The current study revealed the significant contribution of cohesion to L2 writing proficiency beyond the influence of the other linguistic variables, showing that it has a significant and unique contribution to L2 writing

[5] The e-rater (https://www.ets.org/erater.html) utilizes AI technology and natural language processing to assess the essay writings. It offers automatic scoring and feedback, providing such detailed information as grammar, mechanics, word choice, complexity, style, and organization of texts.

© 2024 KASELL All rights reserved

proficiency. However, some limitations exist that should be addressed. Most of all, as aforementioned, all of the essays analyzed in the current study were written within 30 minutes, resulting in relatively short essays. It may have restricted the use of linguistic features across different proficiency levels. In other words, some of the linguistic features may not have been adequately captured, which could have introduced some biases to the findings of the current study. Thus, the findings of the current study should be interpreted with some caution, and more studies should be conducted in diverse contexts to investigate the unique role of cohesion in L2 writing proficiency.

6. Conclusion

The current study investigated the use of cohesive devices and their contribution to L2 writing proficiency by analyzing large-scale writing data from Korean learners of English. It confirmed the previous findings that semantic relatedness (LSA given/new) and the use of connectives (causal connectives, temporal connectives, and logical connectives) are essential variables in predicting L2 writing proficiency. It was also revealed that cohesion had a unique contribution to L2 writing proficiency even after controlling for the influence of lexical and syntactic complexity and n-grams. The research unveiled the distinctive contribution of cohesion to L2 writing proficiency, a contribution that might be obscured by other linguistic features. However, it should be noted that the results of the analysis were largely affected by the length of essays, showing relatively small variances in the use of cohesive devices. Future studies may yield more valid and reliable results by allowing participants a longer time to compose essays and would also benefit from exploring L2 writings across different L1 backgrounds, task types, and genres, revealing the multidimensional aspects of L2 writing and deepening understanding of the role of cohesion in L2 writing.

References

Ädel, A. and B. Erman. 2012. Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes* 31(2), 81-92.

Ai, H. and X. Lu. 2010. A web-based system for automatic measurement of lexical complexity. In *27th Annual Symposium of the Computer-Assisted Language Consortium (CALICO-10)*, 8-2.

Ai, H. and X. Lu. 2013. A corpus-based comparison of syntactic complexity in NNS and NS university students' writing. In A. Díaz-Negrillo, N. Ballier and P. Thompson, eds., *Automatic Treatment and Analysis of Learner Corpus Data*, 249-264. Amsterdam: John Benjamins.

Bestgen, Y. 2017. Beyond single-word measures: L2 writing assessment, lexical richness and formulaic competence. *System* 69, 65-78.

Bestgen, Y. and S. Granger. 2014. Quantifying the development of phraseological competence in L2 English writing: An automated approach. *Journal of Second Language Writing* 26, 28-41.

Biber, D. and B. Gray. 2016. *Grammatical Complexity in Academic English: Linguistic Change in Writing*. Cambridge: Cambridge University Press.

Chon, Y. V., D. Shin and G. E. Kim. 2021. Comparing L2 learners' writing against parallel machine-translated texts: Raters' assessment, linguistic complexity and errors. *System* 96, 102408.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. London: Academic Press.

© 2024 KASELL All rights reserved

Cortes, V. 2013. Connecting lexical bundles and moves in research article introductions. *Journal of English for Academic Purposes* 12(1), 33-43.

Crossley, S. A. 2020. Linguistic features in writing quality and development: An overview. *Journal of Writing Research* 11(3), 415-443.

Crossley, S. A. and D. S. McNamara. 2009. Computational assessment of lexical differences in L1 and L2 writing. *Journal of Second Language Writing* 18(2), 119-135.

Crossley, S. A. and D. S. McNamara. 2012a. Predicting second language writing proficiency: The roles of cohesion and linguistic sophistication. *Journal of Research in Reading* 35(2), 115-135.

Crossley, S. A. and D. S. McNamara. 2012b. Detecting the first language of second language writers using automated indices of cohesion, lexical sophistication, syntactic complexity and conceptual knowledge. In S. Jarvis and S. A. Crossley, eds., *Approaching Language Transfer through Text Classification: Explorations in the Detection-based Approach*, 106-126. Bristol: Multilingual Matters.

Crossley, S. A. and D. S. McNamara. 2013. Applications of text analysis tools for spoken response grading. *Language Learning & Technology* 17, 171-192.

Crossley, S. A., K. Kyle and D. S. McNamara. 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., K. Kyle, L. K. Allen, L. Guo and D. S. McNamara. 2014. Linguistic microfeatures to predict L2 writing proficiency: A case study in automated writing evaluation. *Journal of Writing Assessment* 7(1), 1-34.

Davies, M. 2009. The 385+ million word Corpus of Contemporary American English (1990-2008+): Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics* 14, 159-190.

Douglas, S. R. 2013. The lexical breadth of undergraduate novice level writing competency. *The Canadian Journal of Applied Linguistics* 16(1), 152-170.

Eddington, D. 2015. *Statistics for Linguistics Research*. Cambridge Scholars Publishing.

Garner, J., S. A. Crossley and K. Kyle. 2019. N-gram measures and L2 writing proficiency. *System* 80, 176-187.

Graesser, A. C., D. S. McNamara, M. M. Louwerse and Z. Cai. 2004. Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods* 36(2), 193-202.

Guo, L., S. A. Crossley and D. S. McNamara. 2013. Predicting human judgments of essay quality in both integrated and independent second language writing samples: A comparison study. *Assessing Writing* 18(3), 218-238.

Jung, Y., S. A. Crossley and D. S. McNamara. 2019. Predicting second language writing proficiency in learner texts using computational tools. *Journal of Asia TEFL* 16(1), 37-52.

Kelly, P. C. 1988. *Accessing a Global Language: A Comparison of G-TELP and TOEFL Performance.* Doctoral dissertation, University of San Diego, San Diego, CA, USA.

Kim, H. and H. Seol. 2021. A study on the validity of the G-TELP level test. *The Research Institute of Korean Education* 39, 23-38.

Kim, J. 2014. Predicting L2 writing proficiency using linguistic complexity measures: A corpus-based study. *Korea Association of Teachers of English* 69(4), 27-51.

Kim, J. 2022. The use of cohesive devices in Korean EFL writing across different proficiency levels. *Korean Journal of English Language and Linguistics* 22*,* 1078-1100.

Kim, M. and S. A. Crossley. 2018. Modeling second language writing quality: A structural equation investigation of lexical, syntactic, and cohesive features in source-based and independent writing. *Assessing Writing* 37, 39-56.

© 2024 KASELL All rights reserved

Kim, S. and M. Kessler. 2022. Examining L2 English university students' uses of lexical bundles and their relationship to writing quality. *Assessing Writing* 51, 100589.

Kyle, K. and S. A. Crossley. 2015. Automatically assessing lexical sophistication: Indices, tools, findings, and application. *TESOL Quarterly* 49(4), 757-786.

Kyle, K. and S. A. Crossley. 2016. The relationship between lexical sophistication and independent and source-based writing. *Journal of Second Language Writing* 34, 12-24.

Kyle, K., S. A. Crossley and C. Berger. 2018. The tool for the automatic analysis of lexical sophistication (TAALES): version 2.0. *Behavior Research Methods* 50(3), 1030-1046.

Lee, C., H. Ge and E. Chung. 2021. What linguistic features distinguish and predict L2 writing quality? A study of examination scripts written by adolescent Chinese learners of English in Hong Kong. *System* 97, 102461.

Lee, Y. 2021. The relationship between text cohesion features and English proficiency for Korean college students. *Korean Journal of English Language and Linguistics* 21, 435-449

Lu, X. 2010. Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics* 15(4), 474-496.

Lu, X. 2011. A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly* 45(1), 36-62.

Lu, X. 2012. The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal* 96(2), 190-208.

McNamara, D. S., S. A. Crossley and P. M. McCarthy. 2010. Linguistic features of writing quality. *Written Communication* 27(1), 57-86.

McNamara, D. S., A. C. Graesser, P. M. McCarthy and Z. Cai. 2014. *Automated Evaluation of Text and Discourse with Coh-Metrix.* Cambridge: Cambridge University Press.

Nasseri, M. and P. Thompson. 2021. Lexical density and diversity in dissertation abstracts: Revisiting English L1 vs. L2 text differences. *Assessing Writing* 47, 100511.

O'Leary, J. A. and R. Steinkrauss. 2022. Syntactic and lexical complexity in L2 English academic writing: Development and competition. *Ampersand* 9, 100096.

Read, J. 2000. *Assessing Vocabulary*. Cambridge: Cambridge University Press.

Ryu, J. 2020. Predicting second language writing proficiency in the different genres of writing using computational tools. *Korean Journal of Applied Linguistics* 36(1), 141-170.

Salazar, D. 2014. *Lexical Bundles in Native and Non-Native Scientific Writing: Applying a Corpus-Based Study to Language Teaching*. Amsterdam: John Benjamins.

Stubbs, M. 2007. An example of frequent English phraseology: Distribution, structures and functions. In R. Facchinetti, ed., *Corpus Linguistics 25 Years On*, 87-105. Leiden: Brill.

Tywoniw, R. and S. A. Crossley. 2019. The effect of cohesive features in integrated and independent L2 writing quality and text classification. *Language Education and Assessment* 2(3), 110-134.

Utsunomiya, Y., M. Maruyama and S. Ogasawara. 2016. Estimating TOEIC scores using G-TELP scores: A Bayesian model in a Japanese national university. *JLTA Journal* 19, 27-45.

Yang, W. and Y. Sun. 2012. The use of cohesive devices in argumentative writing by Chinese EFL learners at different proficiency levels. *Linguistics and Education* 23(1), 31-48.

Yoon, H. J. 2021. Challenging the connection between task perceptions and language use in L2 writing: Genre, cognitive task complexity, and linguistic complexity. *Journal of Second Language Writing* 54, 100857.

Zhang, X., X. Lu and W. Li. 2022. Beyond differences: Assessing effects of shared linguistic features on L2 writing quality of two genres. *Applied Linguistics* 43(1), 168-195.

© 2024 KASELL All rights reserved

Zhang, X. and W. Li. 2021. Effects of n-grams on the rated L2 writing quality of expository essays: A conceptual replication and extension. *System* 97, 102437.

Examples in: English
Applicable Languages: English
Applicable Level: All

Appendix 1

Indices of Lexical and Syntactic Complexity (LCA & L2SCA, Lu 2010, 2012)

Lexical Complexity	Indices	Formula
Lexical density	Lexical density (LD)	N_{lex}/N
Lexical sophistication	Lexical Sophistication-I (LS1)	N_{slex}/ N_{lex}
	Lexical Sophistication-II (LS2)	T_s/T
	Verb Sophistication –I (VS1)	T_{sverb}/ N_{verb}
	Corrected VS1 (CVS1)	$T_{sverb}/ \sqrt{2Nverb}$
	Verb Sophistication –II (VS2)	T^2_{sverb}/ N_{verb}
Lexical variation		
NDW	Number of Different Words (NDW)	T
	NDW (first 50words) (NDW-50)	T in the first 50 words of the sample
	NDW (expected random 50) (NDW-ER50)	Mean T of 10 random 50-word sample
	NDW (expected sequence 50) (NDW-ES50)	Mean T of 10 random 50-word sequence
TTR	Type-Token Ratio (TTR)	T/N
	Mean segmental TTR (50) (MSTTR)	Mean TTR of all 50-word segments.
	Corrected TTR (CTTR)	$T/\sqrt{2N}$
	Root TTR (RTTR)	$T/\sqrt{N}$
	Logarithmic TTR (LogTTR)	LogT/LogN
	Uber Index (Uber)	$Log^2N/Log(N/T)$
Verb diversity	Verb Variation-I (VV1)	T_{verb}/ N_{verb}
	Squared VV1 (SVV1)	T^2_{verb}/ N_{verb}
	Corrected VV1 (CVV1)	$T_{verb}/ \sqrt{2Nverb}$
Lexical word diversity	Lexical Word Variation (LV)	T_{lex}/ N_{lex}
	Verb Variation-II (VV2)	T_{verb}/ N_{lex}
	Noun Variation (NV)	T_{noun}/ N_{lex}
	Adjective Variation (AdjV)	T_{adj}/ N_{lex}
	Adverb Variation (AdvV)	T_{adv}/ N_{lex}
	Modifier Variation (ModV)	$(T_{adj} + T_{adv})/ N_{lex}$

© 2024 KASELL All rights reserved

Syntactic Complexity	Indices	Formula
Length of production	Mean length of sentence (MLS)	N of words/ N of sentences
	Mean length of T-unit (MLT)	N of words/ N of T-units
	Mean length of clause (MLC)	N of words/ N of clauses
Sentence complexity	Clauses per sentence (C/S)	N of clauses/ N of sentences
Subordination	Clauses per T-unit (C/T)	N of clauses/ N of T-units
	Complex T-unit per T-unit (CT/T)	N of complex T-units/ N of T-units
	Dependent clauses per clause (DC/C)	N of dependent clauses /N of clauses
	Dependent clauses per T-unit (DC/T)	N of dependent clauses /N of T- units
Coordination	T-units per sentence (T/S)	N of T-units/ N of sentences
	Coordinate phrases per clause (CP/C)	N of coordinate phrases/ N of clauses
	Coordinate phrases per T-unit (CP/T)	N of coordinate phrases/ N of T-units
Phrasal Complexity	Complex nominals per T-unit (CN/T)	N of complex nominals / N of T-units
	Complex nominals per clause (CN/C)	N of complex nominals /N of clauses
	Verb phrases per T-unit (VP/T)	N of verb phrases /N of T-units

Appendix 2

Indices of Academic/ Spoken Bigram and Trigram (TAALES 2.2, Kyle & Crossley, 2015)

Measures	Bigram (19 indices)	Trigram (24 indices)	
Frequency	Raw	Raw	
	Logarithm	Logarithm	
Range	Raw	Raw	
	Logarithm	logarithm	
Association Strength	Bigram MI	Trigram MI	Trigram 2 MI
	Bigram MI^2	Trigram MI^2	Trigram 2 MI^2
	Bigram T	Trigram T	Trigram 2 T
	Bigram ΔP	Trigram ΔP	Trigram 2 ΔP
	Bigram AC	Trigram AC	Trigram 2 AC
Proportion	Proportion 10K	Proportion 10K	
	Proportion 20K	Proportion 20K	
	Proportion 30K	Proportion 30K	
	Proportion 40K	Proportion 40K	
	Proportion 50K	Proportion 50K	
	Proportion 60K	Proportion 60K	
	Proportion 70K	Proportion 70K	
	Proportion 80K	Proportion 80K	
	Proportion 90K	Proportion 90K	
	Proportion 100K	Proportion 100K	

Note 1: TAALES provides five types of association strength of n-grams. Delta P (ΔP) is more sensitive to directionality than other types of association strength. E.g., *most of* and *of most* have the same T-score while having different ΔP.

Note 2: Trigram MI refers to the association strength between the first two words and the last word, and Trigram 2 MI calculates the association strength between the first word and the last two words.

Note 3: Proportion 10k refers to the ratio of the number of n-grams that list in the 10,000 most frequent n-grams in COCA to the numbers n-grams in a text. Proportion is regarded as a variant of frequency.

Note 4: TAALES has 86 n-gram indices, including 43 indices from the academic subsection and the other 43 indices from the spoken subsection of COCA.

© 2024 KASELL All rights reserved

Appendix 3

Descriptive Statistics of Cohesion Indices

Indices	Mean	SD	Minimum	Maximum
local noun overlap	0.42	0.20	0	1
local argument overlap	0.60	0.18	0	1
local stem overlap	0.52	0.21	0	1
global noun overlap	0.35	0.17	0.03	1
global argument overlap	0.52	0.17	0.05	1
global stem overlap	0.44	0.19	0.04	1
local content word overlap	0.13	0.05	0	0.35
global content word overlap	0.10	0.04	0.02	0.39
local anaphor overlap	0.36	0.19	0	1
global anaphor overlap	0.15	0.12	0	1
local LSA overlap	0.22	0.09	0.05	0.55
global LSA overlap	0.20	0.09	0.02	0.61
text LSA overlap	0.27	0.22	0.06	0.85
LSA given/new	0.31	0.05	0.13	0.49
causal connectives	38.06	15.25	0	108.91
contrastive connectives	55.25	16.45	14.56	124.44
temporal connectives	18.78	10.02	0	63.43
additive connectives	15.40	9.61	0	58.82
logical connectives	14.62	8.7	0	55.56
tense and aspect repetition	0.79	0.10	0.33	1

© 2024 KASELL All rights reserved