2020 2022

Assessing English Writing of Japanese Learners: The Development of Scoring
Algorithms of Emails, Descriptions, and Persuasive Essays

Assessing English Writing of Japanese Learners: The Development of Scoring
Algorithms of Emails, Descriptions, and Persuasive Essays

Kim, MInkyung

3,300,000

L2   2
L2
L2

This project investigated longitudinal changes in lexical and syntactic features among
beginning-level EFL learners' writing, and highlighted the lexical and phraseological differences
between L2 written and spoken output.

This project comprised two studies that scrutinized linguistic
characteristics of second language (L2) writing. The first study investigated the longitudinal
patterns of change in lexical and syntactic features in L2 writing. The second study explored
dissimilarities in lexical and phraseological complexity features between L2 written and spoken
opinion responses.

Second Language Writing

Second language writing

The primary objective of this project was initially focused on the creation of machine scoring algorithms capable of accurately predicting the scores assigned by humans to Japanese English as a foreign language (EFL) student writing samples. However, various challenges, such as the unforeseen impact of COVID-19 and the relocation from Japan to South Korea, significantly impeded the execution of the original plan.

This project comprised two comprehensive studies that scrutinized the linguistic characteristics of second language (L2) writing. The primary aim of the first study was to investigate the longitudinal patterns of change in lexical and syntactic features in the writing of novice L2 learners over an extended period. The second study sought to explore dissimilarities in lexical and phraseological complexity features between L2 written and spoken opinion responses.

(1) Research method of the first study: Longitudinal changes in lexical and syntactic features in L2 writing

The first study examined longitudinal changes in lexical and syntactic features in the writing of beginning-level learners in an L2 context.

Data of the first study

The data utilized in the first study were sourced from the Gachon Learner Corpus, which was collected from Korean university students studying English as a foreign language (EFL). Each participant was requested to write ten English essays over a span of ten weeks during an academic semester as part of various English language courses, such as "Academic English" and "Speaking Practice." These courses convened twice a week, with each session lasting 100 minutes, over a period of 14 weeks. To gather the corpus, two sets of ten distinct prompts were employed. The students were instructed to produce essays of 100 to 150 words for each assignment as out-of-class tasks, without any time constraints, and with the possibility of referencing relevant materials. The prompts encompassed topics related to everyday life, including movies, hotel services, and beauty products. The prompts were presented in a fixed order to ensure that participants responded to the same prompt for each writing occasion.

Linguistic measures of the first study

In the first study, a range of linguistic measures were employed to assess the essays.

Lexical complexity was measured in terms of lexical density, variation, and sophistication. Lexical density was measured as the ratio of the number of lexical tokens to the total number of tokens in a text. Lexical variation for content words (CWs) and function words (FWs) was measured using the measure of textual lexical diversity (MTLD, the mean length of sequential word strings in a text that maintain a predetermined TTR value, 0.720). Lexical sophistication for CWs and FWs was measured using word frequency and age of acquisition, which are widely used as proxies for word difficulty.

The measurement of syntactic complexity in the first study involved the utilization of the Syntactic Complexity Analyzer (SCA), which includes a built-in version designed for the automatic analysis of syntactic sophistication and complexity. The SCA draws upon previous studies on L2 writing development and is widely employed in L2 syntactic complexity research. The SCA calculates 14 distinct indices using various linguistic units, such as words, sentences, clauses (defined as structures with a subject and a finite verb), dependent clauses, T-units, coordinate phrases, and complex nominals. From these 14 measures, five were selected for this study based on their construct distinctiveness (targeting different subconstructs of syntactic complexity) and consideration of previous research findings: MLT (Mean Length of T-unit) for assessing global complexity, DC/C (Dependent Clauses per Clause) to evaluate

subordination, MLC (Mean Length of Clause) to measure subclausal complexity, CP/C (Coordinate Phrases per Clause) to analyze coordination, and CN/C (Complex Nominals per Clause) to assess noun phrase complexity.

(2) Research method of the second study: Lexical and phraseological differences between L2 written and spoken opinion responses

The second study examined differences in lexical and phraseological complexity features between L2 written and spoken opinion responses.

Data of the second study

The data utilized in this study consisted of second language (English) writing and speaking performance data obtained from test-takers who participated in the Examination for the Certificate of Competency in English (ECCE), which aligns with the Common European Framework of Reference for Languages (CEFR). The ECCE, developed by Michigan Language Assessment, is designed to evaluate English proficiency at a high-intermediate level, specifically targeting the B2 level of the CEFR. A total of 238 test-takers were included in the analysis, and they provided both writing and speaking samples as part of the ECCE speaking and writing sections. The data comprised paired samples, enabling a comparison of performance across both modalities.

Linguistic measures of the second study

In this study, lexical complexity was assessed based on three measures: lexical density, diversity, and sophistication. Lexical density was computed as the ratio of content word tokens to the total number of tokens present in a text. To evaluate lexical diversity, the hypergeometric distribution diversity index (HD-D) was employed. This index calculates the probability of encountering a token from a specific word type within a random sample of 42 tokens in the text. It provides insights into the diversity of vocabulary usage. Lexical sophistication was determined using two proxies: word frequency and age of acquisition. These proxies are commonly utilized to gauge the difficulty level of words. By considering word frequency and the age at which a word is typically acquired, an estimation of lexical sophistication can be derived.

Phraseological complexity was measured through the analysis of bigram association strength using t- and MI scores. All bigrams, regardless of the parts of speech of the constituent words, were included in the calculation. T- and MI scores were selected for their wide application in L2 studies. However, the study also aimed to explore potential variations in the use of n-grams between L2 writing and speaking performances, an aspect that is less extensively investigated in the literature.

(1) Results of the first study: Longitudinal changes in lexical and syntactic features in L2 writing

In terms of lexical complexity, the findings of this study indicate that over time, there was a significant decrease in lexical density among beginning-level EFL learners. This suggests that these learners tended to include a higher proportion of function words (FWs) in their texts as they progressed. This change may reflect an increased awareness among beginning-level writers of expressing grammatical relationships by incorporating more FWs, such as articles (e.g., "the" and "a") and prepositions.

Regarding lexical sophistication, there was a significant increase in age-of-acquisition scores for content words (CWs) over time. This suggests that as beginning-level EFL learners engaged in English writing practice, they tended to employ words acquired at a later age, which are considered more difficult. However, no significant changes were observed in terms of lexical variation and frequency.

With regards to syntactic complexity, apart from complex nominals per clause (CN/C), no significant changes were found in the measures used in this study among beginning-level EFL learners' essays over time. The decreasing pattern in CN/C suggests that these learners tended to use fewer complex nominals per clause as they progressed. It is possible that these beginning-level students were not aware of the appropriate usage of complex nominals in academic writing. However, this finding does not provide a comprehensive explanation for the observed decrease in CN/C. Further examination of other linguistic features, such as lexical features, may be necessary to fully

understand changes in syntactic features.

Overall, this study investigated longitudinal changes in lexical and syntactic features among beginning-level EFL learners' writing. The findings suggest that over time, these learners exhibited increased lexical sophistication by employing more complex content words and using a greater number of function words, including those acquired at an earlier age by native English speakers (e.g., pronouns and articles).

This study has been published in the internationally recognized peer-reviewed journal "System" in the year 2021.

(2) Results of the second study: Lexical and phraseological differences between L2 written and spoken opinion responses

We utilized t-tests to compare lexical and phraseological features between L2 written and spoken output in this study. The findings revealed consistent patterns in terms of lexical complexity features. L2 written responses exhibited greater lexical complexity, indicated by higher levels of lexical density, lexical diversity, and lexical sophistication measures, compared to L2 spoken responses.

Regarding phraseological complexity features, different patterns emerged depending on the specific phraseological measures employed. When assessing bigram mutual information (MI) scores, L2 written responses consistently displayed higher MI scores (which tend to increase with low-frequency word combinations) compared to L2 spoken responses, regardless of whether the measure was based on academic or spoken corpora. This phenomenon occurred because bigrams that received higher MI scores in the academic corpus (e.g., "global warming," "vast majority," and "human beings") also tended to receive higher MI scores in the spoken corpus. Consequently, bigrams comprising low-frequency words generally received higher MI scores, regardless of the corpus used for MI score calculation.

Conversely, when bigram association strength was assessed using t-scores (which tend to increase with high-frequency word combinations), different patterns emerged depending on the chosen corpus. L2 written responses exhibited higher t-scores when calculated using the academic corpus, whereas L2 spoken responses had higher t-scores when calculated using the spoken corpus. These contrasting outcomes can be attributed to variations in bigram t-scores depending on the corpus utilized for calculation. For instance, bigrams such as "such as," "for example," and "number of" received higher t-scores when calculated using the academic corpus, whereas their corresponding t-scores when calculated using the spoken corpus were lower. Conversely, bigrams like "this is," "going to," and "I mean" had higher t-scores when calculated using the spoken corpus, while their corresponding t-scores when calculated using the academic corpus were lower. Consequently, these findings indicate that bigram t-scores may be useful in distinguishing between L2 written and spoken output but may not serve as reliable measures of phraseological complexity, given their modality-sensitive nature.

Overall, this study highlights the lexical and phraseological differences between L2 written and spoken output and emphasizes the significance of employing greater lexical and phraseological complexity in a modality-insensitive manner in L2 opinion-giving responses.

This study has been published in the internationally recognized peer-reviewed journal "Frontiers in Psychology" in the year 2023.

	2	2	2	1

Kim, Minkyung, & Crossley, Scott	14
Lexical and phraseological differences between second language written and spoken opinion responses	2023
Frontiers in Psychology	1-13
DOI 10.3389/fpsyg.2023.1068685	

Minkyung Kim	103
Exploring longitudinal changes in lexical and syntactic features in beginning-level EFL learner writing	2021
System	1, 12
DOI 10.1016/j.system.2021.102680	

0

0

0