English Language & Literature Teaching, *Vol. 32,* No. 1 Spring 2026
https://doi.org/10.35828/etak.2026.32.1.43

A Coh-Metrix Analysis of Argumentative Essays by Hong Kong ESL, Chinese EFL, and Native Writers*

Hyenyeong Chung**
(Konkuk University)
Moongee Jeon***
(Konkuk University)

Chung, Hyenyeong, & Jeon, Moongee. (2026). A Coh-Metrix analysis of argumentative essays by Hong Kong ESL, Chinese EFL, and native writers. *English Language & Literature Teaching, 32*(1), 43-71.

This study investigated how L2 writing differs across English learning contexts by comparing argumentative essays from the ICNALE corpus produced by Hong Kong ESL learners, mainland Chinese EFL learners, and native English speakers. Multi-dimensional analyses of lexical, syntactic, and discourse indices derived from Coh-Metrix revealed distinct characteristics across groups. Both Hong Kong ESL learners and mainland Chinese EFL learners showed a tendency to produce shorter sentences than native speakers and exhibited lower readability scores. Chinese EFL writing was characterized by a greater reliance on basic, concrete vocabulary, reflecting explicit classroom-based instruction within limited input conditions. In contrast, Hong Kong ESL learners demonstrated comparatively native-like syntactic compression and local referential cohesion. They nevertheless retained L1-mediated discourse patterns, as reflected in relatively constrained global referential cohesion and a preference for causal over additive connectives. These findings suggest that distinctions between ESL and EFL contexts are better understood as a continuum of linguistic repertoires related to differences in input frequency and social environments. Pedagogically, the study underscores the need for context-sensitive instruction and suggests that EFL contexts prioritize systematic exposure to diverse discourse types, whereas ESL environments focus on refining implicit knowledge into explicit global strategies to mitigate cross-linguistic influence.

[English learning contexts/L2 writing/ESL learners/EFL learners/Coh-Metrix]

---

* This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2023S1A5C2A02095124).
** First author *** Corresponding author

I. INTRODUCTION

Although English has become a global language, the English learning contexts experienced by individual learners vary substantially. According to the Three Concentric Circles Model (Kachru, 1990), English around the world can be classified into three categories. The Inner Circle refers to contexts in which English is used as a native language, such as the United States and the United Kingdom, encompassing a range of social settings in which English functions as a first language. The Outer Circle, by contrast, includes multilingual contexts in which English serves as one of the official languages, as in Singapore and Pakistan. Finally, the Expanding Circle refers to contexts in which English is learned as a foreign language in countries where it is not a native language, such as Korea, China, and Japan. Kachru's classification has provided an important theoretical framework for explaining the sociolinguistic spread and diversity of English. Kachru emphasized that English used in Asian contexts is not merely imitation of native-speaker English, but rather distinctive forms of linguistic practice that have emerged through processes of cultural adaptation (Kachru, 1990, 1998).

This distinction of English learning environments is meaningful because ESL and EFL contexts provide learners with qualitatively different language experiences. In ESL settings, learners encounter English in everyday institutional and social domains (e.g., schools and government offices) and often use it as a common language with speakers from diverse linguistic backgrounds, thereby promoting meaning-focused language use (Iwai, 2011). In contrast, in EFL contexts, English learning is largely confined to classroom instruction or mediated input and is frequently oriented toward instrumental goals such as examination performance or employment, with interaction typically limited to teachers and peers who share the same L1 (Kim et al., 2019; Zakian et al., 2022). As a result, these environments differ not only in the amount of exposure but also in their qualitative characteristics: whereas ESL (SL) contexts facilitate acquisition through naturalistic, communicative use, EFL (FL) contexts place greater emphasis on conscious learning via explicit instruction and controlled practice (Krashen, 1982). Consistent with this view, prior research has reported more positive effects of SL contexts (e.g., study abroad) on language development than FL contexts (Díaz-Campos, 2004; Llanes & Serrano, 2017; Segalowitz & Freed, 2004; Taguchi, 2008), and suggests that even with identical tasks and materials, learners' linguistic activities and interactional patterns may vary across contexts (Storch & Sato, 2020). Taken together, these findings indicate that ESL and EFL environments are associated with qualitatively different language experiences, which may be reflected in learners' language use beyond mere differences in input quantity. Nevertheless, recent research has highlighted the role of digital environments in second language exposure, suggesting that learners in traditionally defined EFL contexts may encounter substantial English input through digital media and

online communities, thereby blurring the traditional ESL/EFL distinction (Peters & Webb, 2018; Sundqvist & Sylvén, 2016). However, in contexts such as Hong Kong, where English occupies an institutionalized role in education, administration, and professional communication, learners may encounter English not only as input but also as a medium of participation in everyday social practices. Such contextual conditions suggest that examining writing patterns across ESL and EFL contexts remains a meaningful way to explore how different learning environments are associated with variations in language use.

In this regard, this study investigates how differences in learning contexts affect ESL and EFL learners' writing. Writing is a mode of language production through which the organization and restructuring of language input in learners' internal linguistic systems can be reliably observed (Williams, 2012). In particular, because the writing process involves less pressure for immediate interaction, learners are more likely to draw consciously on their accumulated language resources and knowledge. Accordingly, writing can reveal differences in language experience across distinct learning contexts. For example, Hyland (2002) argued that an essential pragmatic aspect of writing lies in writers' ability to construct and present an authorial identity aligned with the norms of their discourse communities. His findings on Hong Kong ESL writing further suggest that L2 writers may face constraints in constructing a credible authorial identity. Such findings indicate that writing reflects not only linguistic proficiency but also the cultural conventions and learning contexts in which writers have been immersed.

To explore this issue, the present study examines English writing produced by learners in Hong Kong and mainland China. In Hong Kong, English has been institutionalized as an official language of administration, law, and education under British rule. It continues to function as a primary means of communication in public domains, thereby characterizing Hong Kong as an ESL (Outer Circle) context. While Cantonese remains the primary lingua franca in everyday communication, English is institutionally embedded as a functional requirement for professional and academic participation in key domains (Chan, 2026). Despite this evolving sociolinguistic complexity, the institutional role of English continues to distinguish Hong Kong from the more classroom-centered EFL environment of mainland China. In contrast, in mainland China, English is primarily used as a classroom-based foreign language, and English education is characterized by a test-centered orientation that prioritizes improvements in examination performance (Zhao, 2023). Consistent with this educational context, Chinese EFL learners tend to perceive English learning as a means of meeting external expectations, such as academic achievement and examination success (Zhang & Kim, 2013). Therefore, a comparison between Hong Kong ESL learners and mainland Chinese EFL learners provides an appropriate context for investigating the effects of language learning contexts on English usage while controlling for shared cultural backgrounds.

Furthermore, the present study examines how the English essays produced by learners in different language learning contexts differ from those of native speakers. This approach is grounded in theoretical discussions suggesting that when second language development reaches an advanced level, learners' language use may exhibit structural and functional similarities to that of native speakers (Birdsong, 1992, 2004). By comparing the writing patterns of ESL/EFL learners and native speakers, this study aims to provide empirical evidence highlighting the importance of instructional design that takes learners' linguistic backgrounds and learning contexts into account. Ultimately, the findings underscore the need to move beyond uniform evaluation standards in writing assessment and toward an expanded evaluative perspective that acknowledges linguistic characteristics shaped by differing language learning contexts.

To systematically examine these differences, the present study employed Coh-Metrix, an automated text analysis tool for the quantitative assessment of the linguistic and discourse features of English essays. Writing is a multidimensional process that reflects the interaction of lexical choices, syntactic construction, and discourse organization (Crossley, 2020). Coh-Metrix provides a wide range of indices related to multiple linguistic levels of text, enabling objective comparisons of both local and global properties of texts (Crossley & McNamara, 2011; McNamara et al., 2014). Based on these indices, this study analyzes how English essays produced in different language learning contexts vary across multiple linguistic dimensions.

II. PREVIOUS STUDIES

1. Writing Characteristics of Chinese English Learners

This study aims to compare the English writing of Hong Kong ESL learners and mainland Chinese EFL learners with reference to native speakers. Although these two groups differ in their English learning contexts, they share a common linguistic and cultural background as members of the Chinese cultural sphere. Accordingly, this section first reviews previous studies that have examined the writing characteristics of Chinese English learners, providing a foundation for the subsequent comparative analysis.

The writing characteristics of Chinese EFL learners have been examined across a range of genres that differ in communicative purposes and discourse structures. Among these studies, Yang and He (2025) employed large-scale corpus data to compare lexical richness in English essays written by Chinese university students and native speakers. Their analysis revealed no significant difference in lexical density between the two groups. This result was interpreted as reflecting a possible transfer of linguistic features from Chinese, a language

characterized by a relatively high proportion of content words. In contrast, Chinese EFL learners exhibited significantly lower levels of lexical sophistication and lexical variation than native speakers. The authors interpreted these findings as reflecting constraints on the depth and breadth of lexical use among Chinese EFL learners, which may be attributable to instructional practices that prioritize grammatical accuracy and to limited exposure to English. In a similar vein, Zhang et al. (2021) observed limited lexical richness in Chinese EFL writing, noting a heavy reliance on high-frequency and function words compared to native speakers. Additionally, Lei and Yang (2020) found that while Chinese PhD candidates achieved lexical density comparable to native experts, their lexical diversity remained at a novice level. This finding indicates that professional knowledge can drive information density, even when lexical breadth and flexibility remain constrained.

Several studies have examined syntactic differences between essays written by Chinese EFL learners and native speakers. Ai and Lu (2013) showed that Chinese EFL learners generally exhibit lower levels of syntactic complexity than native speakers, particularly in terms of subordination and phrasal sophistication, relying more on simpler sentence structures and less frequent use of complex nominals. However, Shen et al. (2023) demonstrated that such differences vary depending on the indices examined: while no clear group differences emerged at the clausal level, phrasal-level analyses revealed that Chinese learners used relative clause modifiers and adverbial modifiers less frequently than native speakers. In addition, Shen et al. (2023) reported that earlier-stage learners relied more heavily on nouns as modifiers, interpreting this pattern as evidence of L1 transfer from the modifier–head word order of Chinese.

At the discourse level, Chinese EFL learners also exhibit distinct characteristics compared to native speakers, and such differences have been reported across various writing genres, particularly with respect to the use of connectives. Gao (2016) examined the use of linking adverbials in academic writing by Chinese EFL learners and native speakers and found no significant difference in overall usage density; however, Chinese learners tended to underuse additive and adversative adverbials. Lee (2003), in a comparison of narrative writing, reported that Chinese writers relied more heavily on explicit discourse markers than native speakers and placed greater emphasis on sequencing events in temporal order. This pattern was interpreted as reflecting cultural rhetorical conventions in Chinese writing, which prioritize temporal flow and moral or didactic messages over detailed background elaboration. In addition, Mo (2015) compared the use of causal connectives in argumentative writing produced by Chinese EFL learners and native speakers. The study showed that although Chinese learners employed a more limited range of causal connectives than native speakers, they preferred to express causal relations explicitly, as evidenced by the higher density of causal connectives in their writing.

In summary, previous studies have shown that the writing of Chinese L1 learners of

English is characterized by distinctive features across multiple textual dimensions, including lexical aspects (lower lexical diversity and sophistication), syntactic aspects (differences in the use of subordination and phrasal compression strategies), and discourse-level aspects (functional biases in connective use and rhetorical conventions), when compared with native-speaker writing. This body of research provides a basis for the present study, which compares the writing characteristics of Hong Kong ESL learners and mainland Chinese EFL learners, as both groups share a Chinese-based linguistic and cultural background.

Evidence on Hong Kong ESL writing remains relatively limited, but existing work suggests persistent divergences from native-speaker norms, particularly in discourse realization. For example, Hong Kong learners have been reported to favor more formal or phrasal causal expressions (e.g., since, due to) and show a heavy reliance on certain connectors (e.g., moreover), indicating register-sensitive but still learner-like patterns (Bolton et al., 2002). In addition, Hong Kong student writing has been shown to underuse first-person reference compared to expert academic writing, which has been interpreted as reflecting stance-taking constraints and perceptions of authorial authority (Hyland, 2002). Overall, these findings underscore a more systematic comparison of Hong Kong ESL and mainland Chinese EFL writing across multiple textual dimensions.

2. ESL–EFL Comparisons in L2 Writing

Recent research has examined how differences in English learning contexts shape writing among learners with similar linguistic and cultural backgrounds. This line of work suggests that ESL–EFL contrasts may be reflected across multiple linguistic dimensions of written production.

Barrot and Gabinete (2021) examined writing by Asian learners of English within the CAF (complexity, accuracy, and fluency) framework using the ICNALE corpus (Ishikawa, 2023). Learners from ten Asian contexts were classified into ESL and EFL settings and analyzed at the CEFR B1_2 proficiency level. Overall, ESL learners produced texts with greater complexity, accuracy, and length than those of EFL learners. In contrast, EFL learners showed higher numbers of T-units (one independent clause and all of its embedded or subordinate clauses) and clauses per text, suggesting a tendency to organize discourse into more finely segmented clausal units rather than consolidated complex structures. Importantly, substantial variation was also observed within ESL and EFL groups, indicating that writing performance reflects the combined effects of English learning context and country-specific instructional contexts.

Zhang and Kang (2022) directly compared the writing characteristics of Hong Kong ESL learners and mainland Chinese EFL learners. They employed a practical writing task, specifically an invitation letter introducing a cultural exchange program. Based on this task,

texts produced by 32 Hong Kong ESL learners and 30 mainland Chinese EFL learners with comparable English learning experience were analyzed with a focus on lexical and syntactic features. At the lexical level, Hong Kong ESL learners produced information-dense texts with high lexical density, utilizing a wide range of linguistic resources including nouns, verbs, adjectives, adverbs, prepositional phrases, and connectives. In contrast, mainland Chinese EFL learners demonstrated higher lexical diversity, which the authors attributed to the Mainland Chinese English education system's emphasis on explicit vocabulary learning. At the syntactic level, mainland Chinese EFL learners exhibited a more syntax-oriented pattern, with more frequent use of auxiliaries, clausal complements, parataxis, and indirect object constructions. By comparison, Hong Kong ESL learners exhibited a more native-like' semantic dominance' pattern, utilizing adverbial modifiers and adverbial clause modifiers to concisely compress information and prioritize the effective meaning expression. The authors interpreted this contrast as reflecting differences in writing strategies shaped by learning environments, indicating that learning context influences not only accuracy but also sentence flexibility and efficiency of information delivery.

Expanding on Zhang and Kang (2022), Zhang and Li (2025) carried out an analysis of written texts produced by Hong Kong ESL learners and mainland Chinese EFL learners. Using the same writing task as the previous study, they integrated the CAF framework with discourse-level cohesion indices in their analysis. With respect to complexity, mainland Chinese EFL learners demonstrated significantly higher lexical diversity and syntactic complexity than their Hong Kong counterparts, while no clear group differences were observed in lexical density. In terms of accuracy, Hong Kong learners outperformed mainland learners in lexical accuracy, whereas mainland learners showed higher levels of syntactic accuracy. Regarding fluency, Hong Kong learners obtained significantly higher scores. Notably, the results of cohesion indices indicated that Hong Kong learners exhibited higher values than mainland learners across all cohesion indices, including the connective ratio, content word overlap, and Latent Semantic Analysis (LSA). Although statistical significance was observed only for the content word overlap measure, these findings suggest that Hong Kong learners tend to produce texts with greater local cohesion. Taken together, the results reaffirm that L2 writing patterns can differ substantially depending on learners' English learning contexts and educational contexts, even when they share a common cultural background. Importantly, this study contributes to the literature by extending the analytical scope beyond lexical and syntactic dimensions to include discourse-level indices, thereby offering a more integrated account of how learning contexts shape linguistic characteristics in L2 writing.

In summary, although Hong Kong ESL learners and mainland Chinese EFL learners have been shown to exhibit distinct writing characteristics across multiple textual dimensions, important gaps remain in the literature. Previous research has not fully disentangled the

effects of English learning contexts from country-specific educational and sociocultural factors, nor has it systematically examined how ESL and EFL learners' writing aligns with or diverges from native-speaker norms. To address these gaps, the present study compares the writing patterns of Hong Kong ESL learners and mainland Chinese EFL learners with native-speaker reference data, providing a more integrated account of how learning contexts shape writing across lexical, syntactic, and discourse dimensions.

III. METHODOLOGY

1. Materials

The present study draws on essays from the ICNALE (International Corpus Network of Asian Learners of English) corpus (Ishikawa, 2023), which consists of spoken and written texts produced by Asian learners of English and native speakers. All essays were produced under standardized task conditions in ICNALE, with prompts and writing constraints being strictly controlled across groups (Ishikawa, 2023). In accordance with the research objectives, essays written by native speakers (n = 200), Hong Kong ESL learners (n = 100), and Chinese EFL learners (n = 400) were selected, and the texts were screened following the procedures described below.

First, each participant in the ICNALE corpus produced essays on two different topics. To control for topic-related variation (Yang et al., 2015; Yoon, 2021), the analysis was restricted to essays written on the same topic (*Smoking should be completely banned at all the restaurants in the country*).

Second, the ICNALE corpus includes texts produced by English learners across a range of CEFR (Common European Framework of Reference for Languages) proficiency levels (Council of Europe, 2001). The distribution of CEFR proficiency levels for Hong Kong ESL learners and mainland Chinese EFL learners in the initial dataset is presented in Table 1. The distribution is clearly imbalanced across proficiency levels. Accordingly, the analysis was restricted to texts produced by learners at the B1 level, which constituted the largest proportion in both groups.

TABLE 1
Distribution of CEFR Proficiency Levels in the Initial Dataset

Group	Proficiency			Total
	A2	B1	B2	
Hong Kong	1	82	17	100
Chinese	50	337	13	400

Note. Values represent the number of texts. CEFR proficiency labels are available only for learners in ICNALE; native-speaker essays are not CEFR-classified and served as the reference group.

To address the imbalance in sample sizes between B1-level Hong Kong ESL learners and mainland Chinese EFL learners, random sampling was conducted once. Specifically, each text was assigned a random number using Microsoft Excel, and the texts were then sorted in descending order. Based on this procedure, 80 essays were selected from each group for the final analysis. The sample of 80 essays per group was considered adequate for the present analysis. For the Hong Kong ESL group, the selected essays represent nearly the entire available pool of B1-level writers in the corpus (n = 82), thereby closely reflecting the characteristics of this learner group. For the mainland Chinese EFL group, 80 essays were randomly drawn from a larger pool (n = 337), providing a representative subset of the available data.

In summary, the analysis included 80 argumentative essays per group, all written on the same topic by native speakers as well as Hong Kong ESL and mainland Chinese EFL learners at the CEFR B1 level.

2. Measures

All texts were analyzed using Coh-Metrix, an automated English text analysis tool. In the present study, we selected indices that represent lexical, syntactic, and discourse-level characteristics to capture multiple dimensions of text structure in a balanced manner. Because Coh-Metrix indices are organized into categories in which multiple indices collectively reflect shared underlying linguistic properties (Jeon & Lim, 2009), only those indices appropriate for the quantitative analysis of textual characteristics at different linguistic levels were included. All indices analyzed in this study are presented in Table 2.

TABLE 2
Detailed Coh-Metrix Indices by Linguistic Levels

Linguistic level	Category	Indices
Surface	Descriptive features	Total number of words Total number of sentences Average syllables per word Average words per sentence
Lexical	Content words	Noun incidence score Verb incidence score Adjective incidence score Adverb incidence score
	Lexical diversity	Type-token ratio of all content words
	Word frequency	CELEX Log frequency of content words
	Word information	Imageability of content words Concreteness of content words Age of acquisition of content words
	Pronoun	First-person pronoun incidence score Second-person pronoun incidence score Third-person pronoun incidence score

Syntactic	Syntactic complexity	Modifiers per noun phrase Words before main verb
Discourse	Reference cohesion	Argument overlap for adjacent sentences Argument overlap for all sentences
	Connectives	Incidence score of causal connectives Incidence score of additive connectives Incidence score of temporal connectives
Text	Readability	Flesch Reading Ease score Flesch-Kincaid Grade level

Note. The frequency of content words, pronouns, and connectives was calculated per 1,000 words to ensure consistency across the indices.

The indices selected in the present study span five dimensions (surface, lexical, syntactic, discourse, and readability) and together provide an analytical framework for examining how different English learning contexts are reflected in L2 writing across multiple levels. Specifically, lexical and syntactic indices allow comparisons of expressive sophistication and strategies for information compression across learning contexts. Discourse-level indices provide an empirical basis for comparing differences in the acquisition of rhetorical conventions and patterns of organizing information across language contexts. In addition, readability indices are informative for quantitatively comparing the complexity of stylistic choices and communicative strategies adopted by learners under different learning contexts.

3. Statistical Analysis

In the statistical analyses, one-way analyses of variance (ANOVAs) were conducted with the writer group as the independent variable and the various Coh-Metrix measures as dependent variables. All analyses were performed using R version 4.4.3 (R Core Team, 2025), with the significance level set at .05. When a significant main effect of group was observed, post hoc comparisons were carried out using Tukey's Honestly Significant Difference (HSD) test. Both the one-way ANOVAs and post hoc tests were implemented using the base stats package in R (R Core Team, 2025). In addition, effect sizes were calculated using the *effectsize* package (Ben-Shachar et al., 2020), and partial eta squared (η^2_p) values were reported to quantify the magnitude of group differences. All tables in the Results section present between-group analysis results, with descriptive statistics reported as means and standard deviations in parentheses. For reference, partial eta squared values of .01, .06, and .14 are commonly taken to represent small, medium, and large effect sizes, respectively, following Cohen (1988).

IV. RESULTS AND DISCUSSION

1. Surface Level

Surface-level measures capture the basic physical properties of a text, including the number of sentences and words, as well as their length. Table 3 presents the results of these surface-level indices.

TABLE 3
Descriptive Statistics and Results of Analyses for Surface Level Indices

Index	Native	Hong Kong	Chinese	*F*	*p*	η^2_p
Total number of words	224.81 (23.65)	230.24 (32.61)	230.15 (28.63)	0.95	.389	.01
Total number of sentences	9.83 (3.13)	14.11 (3.26)	14.99 (3.25)	59.11	.000	.33
Average syllables per word	1.48 (0.11)	1.53 (0.10)	1.46 (0.07)	11.11	.000	.09
Average words per sentence	24.34 (5.14)	17.02 (4.01)	16.01 (3.83)	86.86	.000	.42

With respect to the quantitative aspects of words and sentences, no significant difference was found in the total number of words across the three groups. In contrast, a significant group difference was observed in the total number of sentences. Post-hoc analyses revealed that essays written by Hong Kong ESL learners and mainland Chinese EFL learners contained significantly more sentences than those written by native speakers (*ps* < .05). Next, analyses of length-related indices showed significant differences among the three groups. Post hoc tests on average syllables per word (i.e., word length) indicated that Hong Kong ESL learners used significantly longer words than both native speakers and mainland Chinese EFL learners (*ps* < .05). By contrast, analyses of average words per sentence (i.e., sentence length) showed that both Hong Kong ESL learners and mainland Chinese EFL learners produced significantly shorter sentences than native speakers (*ps* < .05).

These findings suggest differences in sentence construction patterns between native speakers and Chinese English learners. Specifically, compared to native speakers, Hong Kong ESL learners and mainland Chinese EFL learners showed a tendency to produce a larger number of relatively shorter sentences. While Ai and Lu (2013) documented a preference for simpler sentence structures among mainland Chinese EFL learners, the present results suggest that similar tendencies may also be observed among Hong Kong ESL learners. In addition, Hong Kong ESL learners exhibited a tendency to use relatively longer words than the other two groups. This may reflect Hong Kong ESL learners' preference for relatively more formal expressions compared to native speakers (Bolton et al., 2002), which tend to be morphologically longer or multisyllabic.

2. Lexical Level

1) Content Words

Content words constitute the core semantic components of a sentence. A higher density of content words indicates greater informational richness in a text (Graesser et al., 2004). In addition, the relative distribution of different types of content words provides insight into the characteristics of a text. Specifically, the proportions of nouns and verbs reflect the degree to which a text is information-oriented, whereas the frequency of adjectives and adverbs captures the extent of rhetorical modification adopted by the writer. Table 4 reports the group-level results for the incidence scores of each content word category.

TABLE 4
Descriptive Statistics and Results of Analyses for Content Words Incidence Scores

Index	Native	Hong Kong	Chinese	*F*	*p*	η^2_p
Noun	221.71 (39.87)	255.86 (29.93)	230.02 (25.61)	24.23	.000	.17
Verb	119.19 (17.13)	126.57 (20.34)	113.26 (18.65)	10.11	.000	.08
Adjective	76.65 (20.57)	78.72 (18.73)	75.84 (18.02)	0.48	.617	.00
Adverb	78.39 (20.64)	68.53 (21.11)	78.24 (16.21)	6.76	.001	.05

For noun incidence, the differences among the three groups were statistically significant. Post-hoc analyses revealed that Hong Kong ESL learners exhibited significantly higher noun incidence scores than the other two groups (*ps* < .05). Similarly, post-hoc analyses of verb incidence showed that essays written by Hong Kong ESL learners contained a significantly higher frequency of verbs compared to those of the other two groups (*ps* < .05). In contrast, no significant group differences were observed for adjective incidence. However, significant differences emerged for adverb incidence. Post-hoc analyses indicated that, contrary to the patterns observed for noun and verb incidence, Hong Kong ESL learners used adverbs significantly less frequently than the other two groups (*ps* < .05).

These results indicate that Hong Kong ESL learners show a strong tendency to convey information in a fact-centered manner rather than a rhetorically elaborative one, as evidenced by their higher frequencies of noun and verb use and lower frequency of adverb use compared to the other two groups. This pattern aligns with the semantic dominance pattern identified in Zhang and Kang (2022), characterized by concise information packaging through noun phrase–based constructions. However, unlike the findings of Zhang and Kang (2022), who observed more frequent adverb use among Hong Kong learners, the relatively low adverb incidence in the present study is likely attributable to differences in writing task genre. Whereas prior research focused on practical texts such as invitation letters, the argumentative essay context examined here may be reflected in Hong Kong learners'

reduced use of emotion- and stance-modifying adverbs and a greater reliance on nouns conveying core concepts and event-centered verbs. As a result, a more objective and information-oriented mode of argumentative discourse is reflected in their writing.

Notably, Chinese EFL learners did not show significant differences from native speakers in their patterns of content word use. This result may be associated with the English education system in mainland China, which emphasizes explicit instruction in classroom settings. This pattern suggests that even intermediate-level learners may be able to produce information-oriented texts comparable to those of native speakers. Moreover, this finding is consistent with Yang and He (2025), who argue that the content-word-oriented characteristics of Chinese may facilitate positive L1 transfer, resulting in language use patterns that approximate native-speaker norms. Taken together, the present results could potentially be interpreted in light of explicit instruction and positive L1 transfer.

2) Lexical Diversity

Lexical diversity was measured using the type–token ratio (TTR) calculated for content words, which indicates the extent to which diverse lexical items are used in a text. A type refers to a distinct lexical item, whereas a token refers to each occurrence of a lexical item in the text (Jeon & Lim, 2009). While each type has a value of one, repeated occurrences of the same lexical item increase the token count, resulting in a lower TTR. Thus, a higher TTR value indicates greater lexical diversity (Hong & Sohng, 2019). Table 5 provides a summary of group-level incidence scores for type-token ratio.

TABLE 5
Descriptive Statistics and Results of Analyses for Lexical Diversity

Index	Native	Hong Kong	Chinese	*F*	*p*	η^2_p
Type-token ratio of all content words	.76 (.06)	.69 (.08)	.70 (.07)	20.92	.000	.15

Differences in TTR values among the three groups were statistically significant. Post-hoc analyses revealed that native speakers exhibited significantly higher TTRs than both Hong Kong ESL learners and Chinese EFL learners (*ps* < .05). This finding is consistent with previous studies reporting that Chinese English learners demonstrate lower lexical diversity than native speakers (Lei & Yang, 2020; Yang & He, 2025; Zhang et al., 2021). The results suggest that L2 learners rely on a relatively limited range of lexical resources compared to native speakers.

Contrary to Zhang and Li (2025), no statistically significant difference in lexical diversity was observed between Hong Kong ESL learners and Chinese EFL learners in the present

study. This discrepancy may be attributed to differences in writing task characteristics: whereas Zhang and Li (2025) examined practical writing, the present study focused on argumentative essays, which typically involve a more clearly defined topic and argumentative structure and thus constrain the range of lexical items available to writers. Furthermore, the Chinese linguistic and cultural background shared by both groups may have led to similar patterns of lexical choice and discourse organization in English writing, thereby attenuating group differences in lexical diversity.

3) Word Frequency

Word frequency refers to how often a particular word occurs in general texts and is commonly used as an indicator of lexical ease, with higher-frequency words generally considered easier to process. In natural language, the distribution of word frequency is highly asymmetrical, such that a small number of words occur very frequently, whereas most words appear relatively infrequently (Zipf, 1935). Owing to this non-normal distribution, word frequency values are typically transformed using the natural logarithm to approximate a normal distribution prior to statistical analysis (Graesser et al., 2004). Table 6 displays group-level incidence values for word frequency.

TABLE 6
Descriptive Statistics and Results of Analyses for Word Frequency

Index	Native	Hong Kong	Chinese	*F*	*p*	η^2_p
Log frequency of content words	2.39 (0.17)	2.24 (0.14)	2.39 (0.11)	30.29	.000	.20

The analysis revealed a statistically significant difference in word frequency among the three groups. Post-hoc analyses showed that Hong Kong ESL learners exhibited significantly lower word frequency values than the other two groups (*ps* < .05). This pattern indicates that Hong Kong ESL learners tended to use more lexically demanding (low-frequency) vocabulary. This finding is consistent with the surface-level analysis showing that Hong Kong ESL learners used relatively longer words, which supports the interpretation that they tend to prefer more formal or refined expressions.

By contrast, Chinese EFL learners did not differ significantly from native speakers in terms of word frequency. This pattern may be attributable to the characteristics of EFL learning contexts, in which learners are predominantly exposed to standardized and commonly used vocabulary through textbook-based instruction. Such learning conditions may be reflected in relatively stable lexical choices centered on high-frequency words. Alongside the results of the lexical diversity analysis, this pattern characterizes Chinese EFL learners as relying on the repeated use of more familiar vocabulary.

4) Word Information

Coh-Metrix provides indices of imageability, concreteness, and age of acquisition, which enable inferences about the characteristics of individual lexical items. Imageability refers to the extent to which a word readily evokes a sensory mental image (Paivio, 1971). Concreteness reflects how specific the concept denoted by a word is (Walker & Hulme, 1999). In general, words with higher imageability and concreteness scores are considered easier to process cognitively (Ellis & Beaton, 1993). In contrast, age of acquisition indicates the stage at which a word is typically acquired; higher values represent later-acquired words, thus entailing greater lexical difficulty (Cortese & Khanna, 2008). Table 7 presents the descriptive statistics and ANOVA results for word information across groups.

TABLE 7[1]
Descriptive Statistics and Results of Analyses for Word Information Indices

Index	Native	Hong Kong	Chinese	*F*	*p*	η^2_p
Imageability	394.86 (20.32)	412.04 (21.04)	407.52 (15.01)	17.61	.000	.13
Concreteness	362.18 (21.62)	382.65 (22.11)	374.53 (15.23)	21.46	.000	.15
Age of acquisition	351.18 (34.30)	342.08 (27.78)	332.41 (19.72)	9.04	.000	.07

Statistically significant group differences were observed across all three word information indices. Post-hoc analyses of imageability revealed that native speakers exhibited significantly lower imageability scores than both Hong Kong ESL learners and Chinese EFL learners (*ps* < .05). With respect to concreteness, all pairwise group comparisons were significant (*ps* < .05), with concreteness scores highest for Hong Kong ESL learners, followed by Chinese EFL learners and native speakers. For age of acquisition, a statistically significant difference was found only between native speakers and Chinese EFL learners (*p* < .05), indicating that Chinese EFL learners used words acquired at an earlier age than native speakers.

Taken together, these results suggest that, relative to native speakers, both Hong Kong ESL learners and Chinese EFL learners tend to use words with more intuitive and concrete meanings. In particular, the significantly lower age of acquisition score observed for Chinese EFL learners indicates a greater reliance on basic, everyday vocabulary typically acquired at earlier stages of language learning.

These results are consistent with previous findings reporting lower lexical sophistication among Chinese learners of English compared to native speakers (Yang & He, 2025; Zhang

[1] A correlation analysis revealed that word frequency was negatively correlated with imageability and concreteness across all groups (Imageability: Native *r* = -.61; Hong Kong *r* = -.47; Chinese *r* = -.34 / Concreteness: Native *r* = -.61; Hong Kong *r* = -.58; Chinese *r* = -.39). These results indicate that word information measures reflect different lexical properties rather than redundant lexical dimensions.

et al., 2021). However, the observation that both learners used words with higher imageability and concreteness than native speakers allows for different interpretations as a function of learners' English learning contexts. For Chinese EFL learners, the tendency to rely on more basic vocabulary can be related to limited opportunities to acquire higher-level lexical items under textbook-centered instructional conditions. In contrast, the higher imageability and concreteness observed in Hong Kong ESL learners' lexical choices may be better understood as reflecting their semantic-dominant writing pattern. As demonstrated in the word frequency analysis, Hong Kong ESL learners possess sufficient lexical knowledge to employ low-frequency words. Nevertheless, as reported by Zhang and Kang (2022), these learners tend to prioritize efficient meaning delivery. From this perspective, Hong Kong ESL learners may select concrete yet communicatively effective lexical items, rather than increasing their use of abstract nouns. This pattern may further be explained by the sociolinguistic context of Hong Kong, where English functions as a practical medium of everyday communication, which is aligned with a preference for lexical choices that facilitate intuitive and efficient information transmission.

5) Pronoun

Pronouns constitute a lexical category that functions to replace noun phrases previously introduced in discourse. Patterns of pronoun use can serve as indicators of speakers' discourse strategies (Chung & Jeon, 2025; Hong, 2018; Hyland, 2002). For example, whether speakers predominantly use first-person or second-person pronouns reflects speaker-oriented versus listener-oriented discourse perspectives. Table 8 summarizes group-level pronoun incidence scores along with the corresponding ANOVA results.

TABLE 8
Descriptive Statistics and Results of Analyses for Pronouns Incidence Scores

Index	Native	Hong Kong	Chinese	*F*	*p*	η^2_p
First-person	14.71 (10.17)	7.68 (5.90)	14.39 (7.29)	19.74	.000	.14
Second-person	2.31 (4.63)	2.49 (5.86)	3.62 (5.88)	1.34	.263	.01
Third-person	10.97 (7.73)	8.92 (6.52)	8.78 (4.74)	2.90	.057	.02

Significant group differences were observed only in the incidence of first-person pronouns. Post-hoc analyses showed that Hong Kong ESL learners used first-person pronouns significantly less frequently than native speakers and Chinese EFL learners (*ps* < .05). This finding suggests that Hong Kong ESL learners tend to avoid explicit self-reference even when constructing argumentative essays.

This pattern is consistent with Hyland's (2002) observation that the avoidance of first-person pronouns by Hong Kong ESL learners reflects discourse-level pressures related to

personal commitment. From another perspective, this tendency may also be considered in relation to the distinctive characteristics of Hong Kong English. English in Hong Kong has historically been influenced by British English. Previous research has noted stylistic differences between British and American English, with American English often showing greater writer visibility using first-person pronouns (Ädel, 2006; Leech et al., 2009). The relatively lower use of first-person pronouns observed among Hong Kong ESL learners may therefore be broadly interpreted in light of stylistic traditions associated with British English. More broadly, these findings may reflect the influence of regional linguistic traditions on patterns of self-reference in writing.

3. Syntactic Level

Two Coh-Metrix indices commonly used to assess the syntactic complexity of a text are Modifiers per noun phrase and Words before main verb. Modifiers per noun phrase refers to the proportion of lexical items that modify the head noun within a noun phrase. Words before main verb denotes the number of words that precede the main verb; as the number of constituents appearing before the main verb increases, readers must maintain more information in working memory to identify the subject. Accordingly, higher values on both indices indicate greater syntactic complexity (Graesser et al., 2004). Table 9 reports group-level syntactic complexity indices together with ANOVA results.

TABLE 9
Descriptive Statistics and Results of Analyses for Syntactic Complexity Indices

Index	Native	Hong Kong	Chinese	*F*	*p*	η^2_p
Modifiers per noun phrase	.71 (.17)	.74 (.16)	.68 (.13)	3.28	.040	.03
Words before main verb	3.89 (1.77)	4.33 (1.53)	4.28 (1.39)	1.89	.153	.02

The modifiers per noun phrase index showed a statistically significant group difference. Post-hoc analyses indicated that only the difference between Hong Kong ESL learners and Chinese EFL learners reached statistical significance ($p < .05$), with Hong Kong ESL learners producing more syntactically complex writing. In contrast, no significant group differences were observed for the words before main verb index. One possible explanation for the absence of group differences in the words before main verb index may relate to the proficiency level of the learners. Given that all learner groups were classified at the CEFR B1 level, it is possible that their ability to produce complex clausal structures is relatively limited. As a result, syntactic complexity differences may emerge more clearly in phrasal-level constructions, such as noun phrase modification, rather than in clause-level structures.

These results do not fully align with previous studies reporting that Chinese EFL learners produce more syntactically complex sentences than Hong Kong ESL learners (Zhang & Kang, 2022; Zhang & Li, 2025). However, it is important to note that the indices used to measure syntactic complexity differ between the previous studies and the present study. Zhang and Kang (2022) focused on differences in the typological use of sentence structures, whereas Zhang and Li (2025) employed sentence length as the primary indicator of syntactic complexity. In contrast, the index that showed a statistically significant difference in the present study reflects the density of internal noun phrase modification, capturing the degree of syntactic compression within a given structure rather than increases in sentence length or clausal embedding. From this perspective, the high density of noun phrase modification observed in Hong Kong ESL learners' writing can be interpreted as a more refined manifestation of the "semantic dominance" and compressive writing characteristic observed by Zhang and Kang (2022) at the syntactic level. These findings also align with Ai and Lu (2013) and Shen et al. (2023), who observed that Chinese EFL learners often lack the phrasal sophistication—such as complex nominals—typical of native discourse. In contrast, Hong Kong ESL learners demonstrated greater use of syntactic compression, transitioning toward a more informationally efficient, phrasal-based stage. Ultimately, these results demonstrate that learning contexts shape the qualitative organization of information, extending beyond the mere accumulation of grammatical knowledge.

4. Discourse Level

1) Referential Cohesion

Referential cohesion refers to the extent to which a noun phrase refers to other elements in adjacent sentences or elsewhere in a text (Graesser et al., 2007). It is assessed by the degree of argument overlap across adjacent sentences and across the text as a whole; higher levels of argument overlap are associated with greater referential links and topical unity across sentences. Table 10 provides an overview of argument overlap indices across groups and their ANOVA results.

TABLE 10
Descriptive Statistics and Results of Analyses for Argument Overlap Indices

Index	Native	Hong Kong	Chinese	F	p	η^2_p
Adjacent sentences	.63 (.22)	.61 (.20)	.53 (.16)	5.15	.006	.04
All sentences	.59 (.19)	.52 (.18)	.48 (.14)	8.55	.000	.07

Results for argument overlap across adjacent sentences revealed a statistically significant difference among the three groups. Post-hoc analyses indicated that Chinese EFL learners

showed significantly lower argument overlap than the other two groups (*ps* < .05). In addition, significant group differences were also observed for argument overlap for all sentences. Post-hoc comparisons showed that native speakers exhibited higher scores than both learner groups (*ps* < .05). Taken together, these results indicate that Chinese EFL learners demonstrate lower referential cohesion at both the local and global levels compared to native speakers. In contrast, Hong Kong ESL learners maintain a level of local referential cohesion comparable to that of native speakers, while exhibiting lower global referential cohesion.

Given that discourse cohesion represents a high-order linguistic capacity to integrate semantic relations at a macro-textual level (Kintsch, 1988), the observed differences in referential cohesion reflect varying degrees of learners' ability to organize discourse beyond the sentence level (Graesser et al., 2014). Consistent with the findings of Zhang and Li (2025), the present results show that Hong Kong ESL learners exhibit relatively higher referential cohesion than Chinese EFL learners. Notably, referential cohesion differed between the two learner groups depending on the level of analysis. Specifically, Hong Kong ESL learners demonstrated a level of local referential cohesion comparable to that of native speakers, whereas their global referential cohesion remained relatively lower. This local–global divergence may be closely related to the characteristics of the English learning context experienced by Hong Kong ESL learners. By using English in daily and educational contexts, these learners appear to develop proficiency in maintaining sentence-level meaning connections and referential links. Such patterns can be interpreted as the natural development of discourse skills required for immediate meaning construction in real-time communicative contexts. In contrast, global referential cohesion across an entire text requires planned discourse organization, a capacity that is typically strengthened through explicit writing instruction or academic writing training. Accordingly, the relatively lower level of global referential cohesion observed among Hong Kong ESL learners suggests that patterns of English use in their learning context may support the development of local discourse cohesion but do not automatically guarantee the development of global referential organization.

2) Connectives

Connectives link sentence constituents or sentences and serve as discourse cues that guide readers in integrating inter-sentential meaning relations. By explicitly signaling logical relationships, they organize the overall flow of meaning and contribute to discourse-level coherence (Graesser et al., 2004). Table 11 presents the descriptive statistics and ANOVA results for connective incidence scores.

TABLE 11
Descriptive Statistics and Results of Analyses for Connectives Incidence Scores

Index	Native	Hong Kong	Chinese	*F*	*p*	η^2_p
Causal	30.53 (12.81)	35.88 (15.05)	34.30 (12.60)	3.30	.039	.03
Additive	12.89 (8.03)	9.46 (6.30)	11.67 (6.37)	5.02	.007	.04
Temporal	14.03 (8.83)	13.41 (10.49)	15.33 (8.29)	0.90	.407	.00

The results indicated statistically significant group differences for causal and additive connectives, but not for temporal connectives. Post-hoc comparisons revealed that for causal connectives, only the difference between native speakers and Hong Kong ESL learners was significant ($p < .05$), with Hong Kong ESL learners using causal connectives more frequently. In contrast, a reversed trend was observed for additive connectives: Hong Kong ESL learners utilized fewer additive connectives than native speakers, although this difference reached statistical significance only in the comparison between these two groups ($p < .05$).

These results can be summarized as follows: Chinese EFL learners exhibited connective usage patterns largely comparable to those of native speakers. However, Hong Kong ESL learners showed a stronger preference for explicitly marking causal relationships than native speakers, while being less inclined to connect sentences in a parallel manner through additive connectives.

These findings provide new evidence that the higher connective ratio observed for Hong Kong ESL learners relative to Chinese EFL learners in Zhang and Li (2025) cannot be generalized across connective types but is specifically confined to causal connectives. Remarkably, when contrasted with native speakers, the connective usage pattern of Hong Kong ESL learners closely aligns with patterns previously reported for Chinese EFL learners. Specifically, Gao (2016) demonstrated that Chinese EFL learners tend to underuse additive connectives, while Mo (2015) reported that they employ causal connectives at a higher density than native speakers.

Taken together, this pattern suggests that, despite their substantial daily use of English, Hong Kong ESL learners remain strongly influenced by Chinese discourse conventions in their use of specific discourse devices. This pattern indicates that the multilingual context in Hong Kong does not simply represent the additive accumulation of knowledge from two languages. Rather, it reflects a distinctive developmental trajectory shaped by complex interactions between cross-linguistic influence and language-specific discourse conventions. In particular, the pronounced use of causal connectives observed among Hong Kong ESL learners can be interpreted as the transfer of underlying principles of Chinese discourse organization into English writing. In addition to cross-linguistic influence, pedagogical factors may also contribute to this pattern. Previous research on English instructional materials used in Hong Kong suggests that textbook explanations of additive connectives are often simplified and overlook their pragmatic and stylistic functions in authentic

discourse (Kwan, 2017). Such instructional approaches may be associated with the observed underuse of additive connectives. As a result, the discourse characteristics of Hong Kong ESL learners exhibit a hybrid interlanguage profile, in which native-like syntactic and local cohesion coexist with global organizational patterns that are characteristic of Chinese L2 writing.

5. Text Level

Coh-Metrix evaluates text readability by computing Flesch Reading Ease (FRE) scores and Flesch–Kincaid Grade Level (FKGL) scores. FRE is calculated based on sentence length and word length and ranges from 0 to 100, with higher scores indicating greater ease of reading. In contrast, FKGL converts the same underlying components into an estimate of U.S. grade level, yielding values ranging from 1 to 12. Higher FKGL scores indicate that a higher level of formal education is required to comprehend the text and are therefore interpreted as reflecting greater text difficulty (McNamara et al., 2014). Although both indices are derived from sentence and word length, they are employed here as integrative measures of overall textual readability. Descriptive statistics and ANOVA results for text readability scores are presented in Table 12.

TABLE 12
Descriptive Statistics and Results of Analyses for Text Readability Indices

Index	Native	Hong Kong	Chinese	*F*	*p*	η^2_p
FRE score	57.01 (7.99)	60.46 (9.19)	67.11 (7.04)	32.00	.000	.21
FKGL score	10.89 (1.24)	8.97 (1.69)	7.84 (1.56)	83.74	.000	.41

The results showed that both readability indices differed significantly across the three groups. Post hoc analyses of the FRE scores indicated that all pairwise comparisons were statistically significant (*ps* < .05), with FRE scores increasing in the order of Chinese EFL learners, Hong Kong ESL learners, and native speakers. Similarly, all pairwise comparisons for FKGL scores were significant (*ps* < .05), with FKGL scores highest for native speakers, followed by Hong Kong ESL learners and Chinese EFL learners.

Results from both readability measures consistently indicate that native speaker texts exhibited the highest level of textual complexity, whereas texts produced by Chinese EFL learners showed the lowest level of difficulty. Texts written by Hong Kong ESL learners occupied an intermediate position between the other two groups. These findings suggest that, in line with typical characteristics of second language writing, Chinese EFL learners tended to produce linguistically simplified texts compared to native speakers. It is worth noting that, despite being L2 learners, Hong Kong ESL learners demonstrated a level of textual complexity approaching that of native speakers, a pattern that may be related to their English

learning context. This pattern aligns closely with the earlier findings from the connectives analysis and provides further empirical support for the claim that Hong Kong ESL learners occupy a "bridging" position between native speakers and Chinese EFL learners, potentially reflecting the differences in linguistic input.

V. CONCLUSION

This study examined the effects of differences in English learning contexts on English writing by comparing argumentative essays produced by Hong Kong ESL learners, mainland Chinese EFL learners, and native English speakers within a shared linguistic background. By employing Coh-Metrix to conduct a multidimensional analysis across lexical, syntactic, and discourse levels, this research provided a comprehensive mapping of linguistic characteristics that distinguish these groups. The results can be summarized in three main respects.

First, as a shared characteristic of L2 learners relative to native speakers, both learner groups tended to produce a larger number of relatively shorter sentences and exhibited lower levels of textual readability and complexity. This pattern suggests that second language writing may reflect a tendency to prioritize clarity of meaning and to reduce processing demands. Second, a characteristic specific to Chinese EFL learners was that, while their patterns of content word use and lexical choice were broadly comparable to those of native speakers, they showed a relatively stronger reliance on words with higher imageability and concreteness, as well as on earlier-acquired vocabulary. This pattern suggests that the combination of a classroom-based, explicitly instructed English learning context and the content-word–oriented nature of Chinese may have shaped a writing pattern that is information-focused yet grounded in relatively basic lexical resources. Third, a distinctive feature observed only among Hong Kong ESL learners was that they demonstrated relatively native-like performance in several dimensions, while simultaneously retaining selective influences of Chinese discourse conventions. Their writing was characterized by adopting relatively longer words and a meaning-compressive tendency centered on nouns and verbs, resulting in a fact-centered mode of expression. While Hong Kong ESL learners exhibited levels of local referential cohesion comparable to those of native speakers, their global referential cohesion remained relatively constrained. In addition, their preference for causal connectives and more limited use of additive connectives suggest that Hong Kong ESL writing integrates native-like structural development with discourse-patterning shaped by cross-linguistic influence.

Ultimately, the present study demonstrates that differences in writing characteristics across groups should not be reduced to mere differences in proficiency. Rather, they should be

interpreted as the result of how English learning contexts and educational practices shape the allocation of linguistic resources and the development of writing patterns. Aligned with a usage-based perspective (Ellis, 2002), our findings are consistent with the view that the quantity and quality of linguistic input, as well as the frequency of actual language use, contribute to the internalization of linguistic patterns through repeated experience. In ESL contexts, learners are exposed to a wide range of linguistic structures and discourse patterns on a regular basis, enabling more fluent and flexible language use. In contrast, learners in EFL environments appear to develop writing patterns that rely more heavily on explicit knowledge acquired through form-focused instruction under conditions of limited input.

Furthermore, the present study suggests that the ESL/EFL distinction could be further refined as an analytical framework that captures a continuum of writing practices shaped by differences in language use opportunities, educational practices, and contextual discourse experiences. This perspective is consistent with Blommaert and Backus's (2013) concept of linguistic repertoires, which emphasizes that language ability should be understood as a constellation of layered resources formed within specific social and educational contexts. From this viewpoint, the Hong Kong ESL context is better understood as a multilingual environment where multiple linguistic repertoires interact to shape writing practices, rather than simply as a context that produces higher English proficiency than EFL contexts.

This interpretation calls for evaluating second language writing with attention to learners' linguistic resources and use contexts, rather than treating it as inherently deficient. From a pedagogical perspective, these findings underscore the importance of moving beyond fragmented instruction that treats sentence construction, lexical choice, and discourse cohesion as isolated skills. Instead, writing instruction should adopt an integrated approach that addresses differences in information organization and discourse patterns as they emerge under different learning contexts. Specifically, in EFL contexts, systematic exposure to a wide range of genres and discourse types is needed to compensate for limited input, alongside explicit instruction and practice targeting sentence-to-sentence connections and global discourse organization. For example, corpus-based tasks that expose learners to authentic connective use and contrastive rhetoric activities that highlight organizational differences between L1 and English discourse may help learners develop greater awareness of discourse organization in EFL contexts. In contrast, ESL contexts call for pedagogical interventions that build on learners' implicitly acquired knowledge from rich exposure and refine it into explicit discourse strategies appropriate for academic and argumentative writing. Practical strategies in ESL contexts may include activities such as text analysis that guides learners to examine how information is compressed through complex noun phrases, peer-review sessions focusing on discourse organization and connective choice, and revision tasks encouraging the restructuring of arguments. Finally, such differentiated approaches point to the need for more context-sensitive assessment frameworks that evaluate writing in terms of

how learners organize discourse resources shaped by their learning contexts.

Future research could extend these findings by examining a wider range of genres to determine whether these linguistic patterns are consistent across different rhetorical tasks. Additionally, further research is needed to investigate whether the observed linguistic patterns remain consistent across various proficiency levels to determine the generalizability of these findings. Moreover, future research should operationalize English exposure as a continuous variable measured through indicators such as frequency of use and amount of input, rather than relying solely on the categorical ESL/EFL distinction, in order to more precisely model the relationship between learning environment and writing development. Furthermore, beyond textual analysis, subsequent studies should consider other sociolinguistic factors to better understand the diverse influences on their linguistic choices. In doing so, we can move toward a more holistic view of L2 writers, recognizing their unique linguistic profiles as adaptive responses to their diverse learning experiences.

REFERENCES

Ädel, A. (2006). *Metadiscourse in L1 and L2 English*. John Benjamins Publishing Company. https://doi.org/10.1075/scl.24

Ai, H., & Lu, X. (2013). A corpus-based comparison of syntactic complexity in NNS and NS university students' writing. In Y. Tono, Y. Kawaguchi, & M. Minegishi (Eds.), *Automatic treatment and analysis of learner corpus data* (pp. 249-264). John Benjamins. https://doi.org/10.1075/scl.59.15ai

Barrot, J., & Gabinete, M. K. (2021). Complexity, accuracy, and fluency in the argumentative writing of ESL and EFL learners. *International Review of Applied Linguistics in Language Teaching*, *59*(2), 209–232. https://doi.org/10.1515/iral-2017-0012

Ben-Shachar, M. S., Lüdecke, D., & Makowski, D. (2020). effectsize: Estimation of effect size indices and standardized parameters. *Journal of Open Source Software*, *5*(56), Article 2815. https://doi.org/10.21105/joss.02815

Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language*, *68*(4), 706-755. https://doi.org/10.1353/lan.1992.0035

Birdsong, D. (2004). Second language acquisition and ultimate attainment. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics* (pp. 82-105). Blackwell Publishing. https://doi.org/10.1002/9780470694138.ch30

Blommaert, J., & Backus, A. (2013). Superdiverse repertoires and the individual. In I. de Saint-Georges & J.-J. Weber (Eds.), *Multilingualism and multimodality: Current*

challenges for educational studies (pp. 11-32). SensePublishers. https://doi.org/10.1007/978-94-6209-266-2_2

Bolton, K., Nelson, G., & Hung, J. (2002). A corpus-based study of connectors in student writing: Research from the International Corpus of English in Hong Kong (ICE-HK). *International Journal of Corpus Linguistics*, *7*(2), 165−182. https://doi.org/10.1075/ijcl.7.2.02bol

Chan, J. Y. H. (2026). English and Putonghua varieties in Hong Kong: Language attitudes and identity. *Journal of Multilingual and Multicultural Development*, *47*(1), 581−605. https://doi.org/10.1080/01434632.2024.2398184

Chung, H. N., & Jeon, M. G. (2025). A comparative analysis of linguistic features in the English essays of Korean, Japanese, and Chinese EFL learners using Coh-Metrix. *English Language & Literature Teaching*, *31*(2), 173−199. https://doi.org/10.35828/etak.2025.31.2.173

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203771587

Cortese, M. J., & Khanna, M. M. (2008). Age of acquisition ratings for 3,000 monosyllabic words. *Behavior Research Methods*, *40*(3), 791−794. https://doi.org/10.3758/BRM.40.3.791

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge University Press.

Crossley, S. A. (2020). Linguistic features in writing quality and development: An overview. *Journal of Writing Research*, *11*(3), 415−443. https://doi.org/10.17239/jowr-2020.11.03.01

Crossley, S. A., & McNamara, D. S. (2011). Understanding expert ratings of essay quality: Coh-Metrix analyses of first and second language writing. *International Journal of Continuing Engineering Education and Life Long Learning*, *21*(2-3), 170−191. https://doi.org/10.1504/IJCEELL.2011.040197

Díaz-Campos, M. (2004). Context of learning in the acquisition of Spanish second language phonology. *Studies in Second Language Acquisition*, *26*(2), 249-273. https://doi.org/10.1017/S0272263104262052

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, *24*(2), 143-188. https://doi.org/10.1017/S0272263102002024

Ellis, N. C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*, *43*(4), 559-617. https://doi.org/10.1111/j.1467-1770.1993.tb00627.x

Gao, X. (2016). A cross-disciplinary corpus-based study on English and Chinese native speakers' use of linking adverbials in academic writing. *Journal of English for Academic Purposes*, *24,* 14-28. https://doi.org/10.1016/j.jeap.2016.08.002

Graesser, A. C., Jeon, M., Yan, Y., & Cai, Z. (2007). Discourse cohesion in text and tutorial dialogue. *Information Design Journal*, *15*(3). 199-213. https://doi.org/10.1075/idj.15.3.02gra

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. *The Elementary School Journal*, *115*(2), 210-229. https://doi.org/10.1086/678293

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, *36*(2), 193-202. https://doi.org/10.3758/BF03195564

Hong, S. C. (2018). A corpus-based study of Asian learners' use of demonstrative anaphora. *Studies in Foreign Language Education*, *32*(4). 181-207. https://doi.org/10.16933/sfle.2018.32.4.181

Hong, S. W., & Sohng, H. S. (2019). Exploring the difficulty levels of the reading materials in High School English textbooks with Coh-Metrix for the 1st year students. *English Language & Literature Teaching*, *25*(2), 133-161.

Hyland, K. (2002). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics*, *34*(8), 1091-1112. https://doi.org/10.1016/S0378-2166(02)00035-8

Ishikawa, S. (2023). *The ICNALE guide: An introduction to a learner corpus study on Asian learners' L2 English* (1st ed.). Routledge. https://doi.org/10.4324/9781003252528

Iwai, Y. (2011). The effects of metacognitive reading strategies: Pedagogical implications for EFL/ESL teachers. *Reading*, *11*(2), 150-159.

Jeon, M. G., & Lim, I. J. (2009). A corpus-based analysis of middle school English 1 textbooks with Coh-Metrix. *English Language Teaching*, *21*(4), 265-292. https://doi.org/10.17936/pkelt.2009.21.4.012

Kachru, B. B. (1990). World Englishes and applied linguistics. *World Englishes*, *9*(1), 3-20. https://doi.org/10.1111/j.1467-971X.1990.tb00683.x

Kachru, B. B. (1998). English as an Asian language. *Links & letters*, *5*, 89-108.

Kim, M. S., Choi, D. I., & Kim, T. Y. (2019). South Korean jobseekers' perceptions and (de) motivation to study for standardized English tests in neoliberal corporate labor markets. *The Asian EFL Journal*, *21*(1), 84-109.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163-182. https://doi.org/10.1037/0033-295X.95.2.163

Krashen, S. (1982). *Principles and practice in second language acquisition*. Pergamon Press.

Kwan, Y. H. (2017). The teaching of additive connectives: An evaluation of current English grammar textbooks in Hong Kong. *The Asian Journal of Applied Linguistics*, *4*(2), 161−172.

Lee, M. Y. (2003). Discourse structure and rhetoric of English narratives: Differences between native English and Chinese non-native English writers. *Text & Talk*, *23*(3), 347-368. https://doi.org/10.1515/text.2003.015

Leech, G., Hundt, M., Mair, C., & Smith, N. (2009). *Change in contemporary English: A grammatical study*. Cambridge University Press.

Lei, S., & Yang, R. (2020). Lexical richness in research articles: Corpus-based comparative study among advanced Chinese learners of English, English native beginner students and experts. *Journal of English for Academic Purposes*, *47*, Article 100894. https://doi.org/10.1016/j.jeap.2020.100894

Llanes, À., & Serrano, R. (2017). The effectiveness of classroom instruction 'at home' versus study abroad for learners of English as a foreign language attending primary school, secondary school and university. *The Language Learning Journal*, *45*(4), 434–446. https://doi.org/10.1080/09571736.2014.923033

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press. https://doi.org/10.1017/CBO9780511894664

Mo, J. (2015). A contrastive study of the use of causal connectives by Chinese EFL learners and English native speakers in writing. *Theory and Practice in Language Studies*, *5*(11), 2426-2432. http://doi.org/10.17507/tpls.0511.30

Paivio, A. (1971). *Imagery and verbal processes*. Holt, Rinehart and Winston.

Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, *40*(3), 551−577. https://doi.org/10.1017/S0272263117000407

R Core Team. (2025). *R: A language and environment for statistical computing (Version 4.4.3)* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Segalowitz, N., & Freed, B. F. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, *26*(2), 173-199. https://doi.org/10.1017/S0272263104262027

Shen, C., Guo, J., Shi, P., Qu, S., & Tian, J. (2023). A corpus-based comparison of syntactic complexity in academic writing of L1 and L2 English students across years and disciplines. *PLOS ONE*, *18*(10), Article e0292688. https://doi.org/10.1371/journal.pone.0292688

Storch, N., & Sato, M. (2020). Comparing the same task in ESL vs. EFL learning contexts: An activity theory perspective. *International Journal of Applied Linguistics*, *30*(1), 50-69. https://doi.org/10.1111/ijal.12263

Sundqvist, P., & Sylvén, L. K. (2016). *Extramural English in teaching and learning: From theory and research to practice*. Springer.

Taguchi, N. (2008). The role of learning environment in the development of pragmatic comprehension: A comparison of gains between EFL and ESL learners. *Studies in Second Language Acquisition*, *30*(4), 423-452. https://doi.org/10.1017/S0272263108080716

Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(5), 1256-1271. https://doi.org/10.1037/0278-7393.25.5.1256

Williams, J. (2012). The potential role (s) of writing in second language development. *Journal of Second Language Writing*, *21*(4), 321-331. https://doi.org/10.1016/j.jslw.2012.09.007

Yang, Y., & He, X. (2025). Lexical richness in Chinese university students' EFL writing: a corpus-based comparison. *Humanities and Social Sciences Communications*, *12*(1), 1-11. https://doi.org/10.1057/s41599-025-05560-x

Yang, W., Lu, X., & Weigle, S. C. (2015). Different topics, different discourse: Relationships among writing topic, measures of syntactic complexity, and judgments of writing quality. *Journal of Second Language Writing*, *28*, 53-67. https://doi.org/10.1016/j.jslw.2015.02.002

Yoon, H. J. (2021). Interactions in EFL argumentative writing: Effects of topic, L1 background, and L2 proficiency on interactional metadiscourse. *Reading and Writing*, *34*(3), 705-725. https://doi.org/10.1007/s11145-020-10085-7

Zakian, M., Xodabande, I., Valizadeh, M., & Yousefvand, M. (2022). Out-of-the-classroom learning of English vocabulary by EFL learners: investigating the effectiveness of mobile assisted learning with digital flashcards. *Asian-Pacific Journal of Second and Foreign Language Education*, *7*(1), Article 16. https://doi.org/10.1186/s40862-022-00143-8

Zhang, C., & Kang, S. (2022). A comparative study on lexical and syntactic features of ESL versus EFL learners' writing. *Frontiers in Psychology*, *13*, Article 1002090. https://doi.org/10.3389/fpsyg.2022.1002090

Zhang, C., & Li, R. (2025). A Study on Linguistic Features of Chinese ESL and EFL Learners' Writing: A Learning Context Perspective. *SAGE Open*, *15*(3). https://doi.org/10.1177/21582440251378082

Zhang, H., Chen, M., & Li, X. (2021). Developmental features of lexical richness in English writings by Chinese beginner learners. *Frontiers in Psychology*, *12,* Article 665988. https://doi.org/10.3389/fpsyg.2021.665988

Zhang, Q. M., & Kim, T. Y. (2013). Cross-grade analysis of Chinese students' English learning motivation: A mixed-methods study. *Asia Pacific Education Review*, *14,* 615-627. https://doi.org/10.1007/s12564-013-9288-3

Zhao, H. (2023). The influence of test-oriented teaching on Chinese students' long-term use of English. *International Journal of Education and Humanities*, *6*(2), 123-128.

Zipf, G. K. (1935). *The psycho-biology of language: An introduction to dynamic philology*. Houghton Mifflin.

Examples in: English
Applicable Languages: English
Applicable Levels: Tertiary

Hyenyeong Chung (researcher)
Konkuk Research Institute for Multilingualism and Multiculturalism
Konkuk University
05029 Neungdong-ro, Gwangjin-gu, Seoul, Korea
Email: chn4050@konkuk.ac.kr

Moongee Jeon (professor)
Department of English Language & Literature
Konkuk University
05029 Neungdong-ro, Gwangjin-gu, Seoul, Korea
Email: mjeon1@konkuk.ac.kr

Received in: January 28, 2026
Reviewed in: March 6, 2026
Revised version received in: March 23, 2026