

## Using Small Corpora of Critiques to Set Pedagogical Goals in First Year ESP Business English

Yu-Chi Wang  
(Feng Chia University)  
Richard Hill Davis  
(Feng Chia University)

**Wang, Y. C., Davis, R. H. (2021). Using small corpora of critiques to set pedagogical goals in first year ESP business English. *Asia Pacific Journal of Corpus Research*, 2(2), 17-29.**

The current study explores small corpora of critiques written by Chinese and non-Chinese university students and how strategies used by these writers compare with high-rated L1 students. Data collection includes three small corpora of student writing; 20 student critiques in 2017, 23 student critiques from 2018, and 23 critiques from the online Michigan MICUSP collection at the University of Michigan. The researchers employ *Text Inspector and Lexical Complexity* to identify university students' vocabulary knowledge and awareness of syntactic complexity. In addition, *WMatrix4®* is used to identify and support the comparison of lexical and semantic differences among the three corpora. The findings indicate that gaps between Chinese and non-Chinese writers in the same university classes exist in students' knowledge of grammatical features and interactional metadiscourse. In addition, critiques by Chinese writers are more likely to produce shorter clauses and sentences. In addition, the mean value of complex nominal and coordinate phrases is smaller for Chinese students than for non-Chinese and MICUSP writers. Finally, in terms of lexical bundles, Chinese student writers prefer clausal bundles instead of phrasal bundles, which, according to previous studies, are more often found in texts of skilled writers. The current study's findings suggest incorporating implicit and explicit instruction through the implementation of corpora in language classrooms to advance skills and strategies of all, but particularly of Chinese writers of English.

**Keywords:** Corpus Linguistics, Academic Writing, Business English, Critique Genre, Pedagogy

### 1. Introduction

Research in corpus linguistics has advanced quickly over the last few decades to support and explain empirical studies of language variation. Corpus linguistics can be seen as a methodological approach to collect texts in natural contexts, analyze the data, and identify linguistic patterns to benefit language learning and teaching (Ai & Lu, 2013). An increasing number of recent language research has examined small corpora to understand better patterns in language used in particular settings. The research was used to help inform the development of language pedagogy for English for Academic Purpose (EAP) and English for Specific Purpose (ESP) courses (Chan, 2018).

The study of learner corpora, referring to electronic collections of texts produced by second or foreign language learners, has also gained increasing attention in second language acquisition and foreign language teaching. For instance, comparisons between a learner corpus and a native English speakers' corpus may help provide a comprehensive understanding of differences in proficiency between native and non-native students' written and spoken language. Previous research has shown a positive correlation between language learners' proficiency levels and their use of structural complexity (Ai & Lu, 2013). Nesselhauf (2004) stated that while previous corpus studies have mainly

focused on corpora of native English speakers, it is also vital to study learner corpora to understand better L2 learners' language use, including their difficulties for a particular group of learners.

The purpose of the current study is to profile how Taiwanese and international students used features of interactional metadiscourse, such as hedges or modals, and grammatical features of written discourse in the critiques they wrote for the First-year Business English course. For the present study, we focused on students in the International Business Administration (BIBA) program in a university in central Taiwan. This is an English-only degree that entails a year's study abroad as a graduation requirement. In 2017, based on a series of conferences with Business faculty and an oral survey of former students, we found discrepancies between students' language proficiency and needs in the business and academic field. Accordingly, we redesigned the curriculum to increase the emphasis on writing in the course. In addition, the unique class included both English as Additional Language (EAL) speakers of English and native speakers (L1 English). This causes us to incorporate features of global Business English for second-language writers, such as a greater focus on project-based and collaborative learning (Chan, 2018).

In addition to requiring students to write critiques, we also compare theirs with critiques from MICUSP (Michigan Corpus of Upper-Level Student Papers, <https://eresources.eli.lsa.umich.edu/micusp-corpus-of-written-academic-papers/>) to identify differences between native and non-native English speakers' language use. For this study, we followed Lee and Deakin (2016) in compiling three small, specialized corpora for analysis: 20 critiques from 2017, 23 critiques from 2018, and 23 critiques from a range of fields in the online Michigan MICUSP collection. Our discussion focuses on pedagogical and corpus-related issues, specifically on using corpora and corpus-based applications to expand pedagogy for first-year English as Additional Language (EAL) learners in business fields.

## 2. Literature Review

### 2.1. Corpus-based BE Writing Development: Corpora, Rhetoric, Complexity

Ilie, Nickerson, and Planken (2019, p. 38) trace the changing emphases in Business English (BE) as located within English for Specific Purposes, from emphases on rhetoric, needs analysis, and skills through the 1980s to the shift in the 1990s and "the first decades of the 2000s...to a more genre-analytic approach" (p. 39). They find critical discourse analysis and a rhetorical perspective on organizational discourse necessary in teaching business discourse. Particularly if a multimodal approach is incorporated using a range of technology-enhanced tasks incorporating computer-supported media for speaking, viewing, presenting, and writing.

Bargiela-Chiappini and Zhang (2013, p. 207) explain that BE began as "a materials-led movement rather than a research-led movement," leading them to call for studies within rhetoric, pragmatics, and discourse. Some of the research needed for curriculum development in BE is suggested by Chan (2018), who includes a list of references in the larger area of curricular needs and materials and the sub-areas of BE. They also ground her 2019 study in which she surveys the kinds of writing and other language needs faced by business professionals as they increase in age and ranking. Chan (2018) finds that just as business professionals need to handle multiple genres to master a range of developing situations, instructors must create greater awareness among learners of the need to master a variety of language learning strategies to meet rapidly changing demands opportunities.

Identifying and understanding rhetorical strategies as they occur in a collection of authentic texts is vital if the writer learns how to move among genres and discourse structures, mainly if the writer must use persuasion, which animates many business reports, case studies, and proposals. Using a corpus of fund-raising letters, Biber, Anthony, and Gladkov (2007) list and exemplify the rational appeals of classical Western rhetoric, which compare well with Kirkpatrick and Xu's (2012) analysis

of Chinese rhetoric. Classical rhetoric (Biber et al., 2007) incorporates Logos (i.e., cause/effect, stages in progress, authority other than the writer); the appeals to the credibility with Ethos (i.e., display of first-hand experience or shared interest with the audience); and affective appeals, with Pathos (i.e., strong or emotional language). All three of these components occur in the genre of critique and can affect writer stance to varying degrees. Grabe and Zhang (2016) note that drawing on smaller corpora allows the researcher to incorporate semantic interpretation. Moreover, they add, “discourse analysis for argumentative/persuasive genres can combine the rhetorical appeals (ethos, logos, pathos) with linguistic/structural features” (p. 256), which may be helpful in combination with a genre-based examination of moves within specific types of writing tasks.

Biber, Gray, and Poonpon (2011) proposed a “hypothesized set of developmental stages for complexity features.” These stages, five in all, were described and tabulated by “grammatical structure and grammatical role” (p. 4). For several years, Biber and his colleagues have employed corpora as they emphasized the role played by phrasal complexity as differentiated from studies highlighting clauses or T-units: see, for example, Biber, Gray, and Poonpon (2013), in which they explain that “the unique grammatical characteristic of advanced academic writing is “its reliance on phrasal structures, especially complex phrases with phrasal modifiers” (2013, p. 192). Staples, Egbert, Biber, and Gray (2016) is even more explicit in terms of developmental phases:

Biber et al. (2011) is based on the premise that novice academic writers begin with the clausal complexity features most common in speech and then gradually develop proficiency in the dense use of the phrasal complexity features associated with specialist academic writing (Staples et al., 2016, p.153).

After reviewing the literature and discussing the relative scarcity of comparisons of syntactic complexity between L1 and L2 writing, Ai and Lu (2013), identified significant differences between the writing of native L1 speakers and EAL speakers. Specifically, EALs had significantly shorter mean lengths of clauses, sentences, and T-units, and “a significantly smaller proportion of dependent clauses and complex T-units” as well as “a significantly smaller proportion of complex nominal” (p. 257-258). While EALs present less subordination, there is no critical difference between the two groups for “sentential coordination” (p. 261), where syntactic complexity can become even more important in its rhetorical functions in academic writing.

In examining interactions between writers and readers, previous literature has shown hedges, boosters, self-mentions, and attitude markers as ways for writers to signal their presence and opinion, and native English speakers are more likely to use more hedges and more types of stance markers than EALs (Bax, Nataksuhara, & Waller, 2019). Gardner, Nesi, and Biber (2019) subject the BAWE (British Academic Written English) corpus to multidimensional and situational (within disciplinary groupings) analysis and find four dimensions, the fourth of which – informational density – characterizes academic writing and includes writings such as Business proposals. Density differs between the “compressed procedural information” of the first dimension and the information density in the fourth dimension. In contrast, stance features vary according to whether the writing is reflective or self-evaluating instead of evaluating outside sources (Gardner et al., 2019, p. 672).

Previous literature has shown strong connections between L2 students’ syntactic complexity and language proficiency level in written work. The expanding uses for prepositions in nominal and adverbial phrases are especially noteworthy as phrasal complexity increases in importance in written text. Using Biber’s original multidimensional model for the analysis of text (Biber, 1988), Kim and Nam (2019) reinforce this finding in their recent study of L2 written argument, while Jitpraneechai (2019) notes that native speakers (L1) produced longer and more complex noun phrases incorporating prepositions than Thai EAL student writers. Moreover, in a study of lexical bundles in written argument, Bychkovska and Lee (2017) note that L1 Chinese professionals were more likely to use verb

phrase bundles, preferring clausal to phrasal bundles.

## 2.2. Relational Uses of Language: Hedges in Stance and Metadiscourse

Studies of stance, particularly of the use of metadiscourse in L1 writing, are abundant; however, scholars of corpus linguistics call for studies of a broader range of constructs and a greater focus on their pedagogical application. Bax et al.'s (2019) study of CEFR (Common European Framework of Reference for Languages) scripts provides information on constructs and application. Based on their study of intermediate and advanced second language writing, they see metadiscourse as having two functions: textual, with functions of cohesion and organization, and interpersonal, which can suggest the writer's stance to the reader. They find that as EAL writers advance in proficiency, they use fewer metadiscourse markers within a greater range.

The literature review in the analysis of successful/less successful first- and second-language writers by Lee and Deakin (2016) is well worth quoting as it stands. They identify critical works in interpersonal and, more importantly for the present study's focus, interactional metadiscourse across disciplines, learning contexts, lingua-cultures, and genres, as well as between first language (L1) and second language (EAL) writers of English. They find that successful writing by all writers, regardless of language, presents significantly more hedges. Hedges were critical in comparing critiques/evaluations, argumentative essays, and response papers by first-year writers with works by upper-level students collected in MICUSP (the Michigan Corpus of Upper-Level Student Papers) and published academic writers in COCA (Corpus of Contemporary American English).

Previous research shows that hedges can be a problem for EAL writers. Gherdan (2019) asserts: they are essential to learning, however, because they let writers "submit new arguments...avoid making absolute or categorical statements [and] ...leave room for other voices" (pp. 124-125). Ho and Li (2018) found that first-year writers' test scores on argumentative essays correlated only with interactional markers. Specifically, hedges and markers for attitude (e.g., "understandable") and engagement, typified by the use of inclusive "we." Reviewing persuasive writing, they found positive relationships to test scores as the range of markers increased. In her subsequent analysis of argumentative essays, including critiques, in MICUSP, Aull found that "100% of the argumentative and explanatory papers contained hedges and boosters" (Aull, 2019, p. 280).

The type of text presented to an audience has an impact: Cheng, Lam, and Kong (2018) find that identifying interpersonal meaning in textbooks on workplace communication has been neglected, particularly for building solidarity and shared understandings, in-text types such as emails and applications. As a partial corrective, Schieber and Robles (2019, p. 302) suggest using student-written reflections to identify where they felt unsure of their ability in "delivering positive, informative, negative, and persuasive messages." In the present study, the final assignment we describe stipulated an audience: the following year's incoming BIBA students, who will presumably share many audience characteristics with the current student-writers. Our requiring outside sources to support their summary, evaluation, and final recommendation of an article to be read was designed to deepen the persuasive quality of the writing as well as provide practice with sources and citation skills.

## 3. Methodology

### 3.1. The BIBA Portfolio

This study examined the differences in language use in written texts between native English speakers and first-year students in the BIBA program. The weekly freshman English course meetings for BIBA students are two hours long. The authors have collaboratively taught the courses, and we focus on instructor-based, team-conducted, project-based interaction. In the first semester, the

curriculum includes business-related telephone calls, survey administration, meeting etiquette, formal teamed presentations for video-recording and subsequent discussion, research skills keyed to teamwork, business-centered language acquisition through assigned readings, and oral analysis of business-related articles, stressing critical thinking. We add an electronic portfolio and weekly vocabulary enhancement in the second semester using the Cambridge University business school wordlist (<http://www.cambridgeenglish.org/images/22099-vocabulary-list.pdf>) as well as the Coxhead academic wordlist (<https://www.vocabulary.com/lists/218701>).

The central focus of the second semester is a portfolio-based assessment with writings in different genres and for different audiences, published to the students' electronic portfolios, including emails to alumni, email to a professor, an article summary, a product description, and a final 1,000-word critique. These include Nesi and Gardner's (2018, p. 53) note that a critique typically occurs as "academic paper review; interpretation of results; legislation evaluation; policy evaluation; programme evaluation; project evaluation..." several of which we find typically in business-related writing tasks. Finally, the student-writer recommends an article to the following year's class in the final paper, reinforcing the choice with three related student-chosen sources. It is designed to combine the genre of critique, as it incorporates both summary and evaluation, with an audience-centered recommendation and citation skills. Before final submission, three drafts are written, posted on the class LMS, and assessed, including a solid peer evaluation. In addition to increasing facility with handling sources, the assignment hopes to prepare students for a range of future business writings. Its goal for students is to improve both their ability to summarize with accuracy and their comfort with evaluation, gain awareness of relational uses of language and multiple genres, and enhance lexical and syntactic complexity.

### 3.2 Data Collection and Analysis

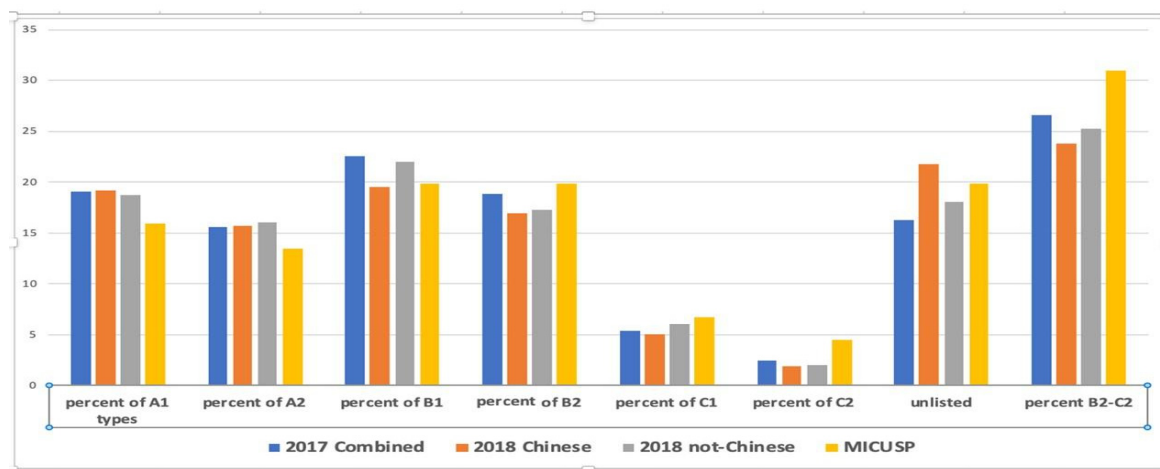
We compiled three small, specialized corpora of BIBA student writings: 20 critiques from 2017, 23 critiques from 2018, and 23 critiques from a range of fields (social sciences and humanities) in the online Michigan MICUSP collection, and submitted each to specialized corpus-analysis tools. To obtain a vocabulary profile and analyze relational uses of language, or writer stance as shown by metadiscourse markers, we used *Text Inspector (TI)* ([textinspector.com](http://textinspector.com)), affiliated with Cambridge University Press. Stephen Bax initially developed the prototype for the program in 2012. It incorporates the English Vocabulary Profile, based on the Cambridge Learner Corpus, and identifies CEFR levels (<https://www.englishprofile.org/wordlists/evp-faqs>). For lexical and syntactic complexity, we used Xiaofei Lu's *Lexical Complexity Analyzer (LCA)* and L2SCA, his *Second Language Syntactic Complexity Analyzer*, both maintained online by Haiyang Ai ([aihaiyang.com/software](http://aihaiyang.com/software)). Lu's *Lexical Complexity Analyzer (LCA)* incorporates 25 metrics used by scholars to analyze lexical density, lexical sophistication, and lexical variation. In addition, we used WMatrix4®, the corpus analysis and management tool online at the University of Lancaster (<https://ucrel.wmatrix4.lancaster.ac.uk/>).

## 4. Research Findings and Discussion

### 4.1. Lexical Richness

After reviewing the variance of reports on validity for both TOEFL and IELTS as entrance tests, Carlsen (2018) noted that there is probably a "minimum entrance requirement to prevent failure caused by language limitations" (p. 78). A variety of statistical analyses of academic mastery by L2 students led her to conclude that vocabulary at the level of CEFR B2 should be the admissions cutoff point. Bax et al. (2019) based their analysis of metadiscourse markers in L2 writing exclusively on Cambridge English General English examinations at CEFR B2, C1, and C2 levels, using *Text Inspector*

(TI). Accordingly, we first analyzed the vocabulary of the first ten-thousand words in each small corpus with TI, as graphed below in Figure 1.



**Figure 1.** EVP Analysis of CEFR-levels in Vocabulary for Four Cohorts

As expected, the graduate student writers of MICUSP critiques had lower percentages of tokens for A1 and A2 words, somewhat higher for B2-C1 and unlisted words, and higher still for C2 and the combined B2-C2 scales for words. Regarding the differences between Chinese L1 and non-Chinese students, the latter group shows a slightly higher percentage of tokens for A2-C1 than the Chinese L1 students. Unlisted words for each of them fell into three categories: proper names of persons, places, and business or technical company names and terms; internet/media-related words and terms; and lemma not included in the CEFR listings, often because of spelling differences between British and American spellings, e.g., *analyze* vs. *analyse*.

In addition, we employed Lu's (2012) *Lexical Complexity Analyzer (LCA)* to analyze lexical density, lexical sophistication, and lexical variation. In Lu's (2012) discussion of these metrics, he chose not to review error analysis. Therefore, of the 25 metrics which he studied and measured against 12 groups of Chinese L2 writers, we report here on only five across 1,000-word samples from our cohorts: 2017 BIBA, 2018 BIBA (Combined, L1 Chinese only, non-Chinese) and MICUSP, because they are the five Lu noted as important. Since Lu found "no evidence of a significant correlation between lexical density and test takers' rankings" (Lu, 2012, p. 198), we chose not to report on lexical density. In addition, and for similar reasons, we decided to bypass reporting the metrics for lexical and verbal sophistication and most of the additional metrics for variation. We do, however, report scores for the number of different words, the corrected and root type-token ratios, and the squared and corrected verb variation metrics because Lu argued they have the most potent effect on test score rankings (Lu, 2012).

**Table 1.** Selected Categories from Lexical Complexity Analysis of Cohort Vocabulary Samples

Lexical Complexity Using <i>LCA</i> : (Ai 2010) [Samples are Second 1000 Words of Full Text]	Number of Different Words	Corrected Type-token Ratio	Root Type-token Ratio	Squared Verb Variation	Corrected Verb Variation
2017: 1000-word Sample	363	8.31	11.76	55.60	5.27
2018: 1000-word Sample	351	7.85	11.10	46.57	4.83
2018: Chinese 1000-word Sample	<b>327</b>	<b>7.41</b>	<b>10.48</b>	<b>28.55</b>	<b>3.78</b>
2018: Non-Chinese 1000-word Sample	405	8.93	12.63	60.43	5.50
Critiques 1000-word Sample	384	8.54	12.08	49.42	4.97

Our focus is on the number of different words and the visible differentiation between Chinese L1 and international (non-Chinese L1) students. Concerned that the selection of samples could have skewed

the analysis, we ran several others. Regardless of which instance we selected, the differentiation pattern remained similar to the one we report here. LCA results show that among the five text samples, Chinese L1 texts have the lowest values in terms of the number of different words, type-token, and verb variation among the three corpora (Table 1). Accordingly, we ran an analysis of the full texts of all final papers for the course using WMatrix4®, which indicated that students with any variety of Chinese as their first language used slightly fewer word types and repeated them more often. The data is shown in Table 2.

**Table 2.** Use of Word Types by Chinese L1 and Non-Chinese L1 Writers

First Language	Number of Types Shown	Total Frequency of Type Shown	Types/Frequency
Chinese	3262	16329 (100%)	19.976 (20%)
Non-Chinese	2156	9587 (100%)	22.488 (23%)

WMatrix4® identifies the difference between type (=lemma) and token (the word in whatever form). The results in Table 2 show that Chinese students wrote more words (tokens) with more word types than non-Chinese students. There were six more Chinese students in the class and thus six more papers with roughly 6000 more words. However, the percentage of types/frequency of tokens is lower. That tells us that Chinese students used proportionally fewer word types, or 20%, compared with non-Chinese, who achieved 23%.

Wmatrix4®'s built-in Log-Likelihood Calculator showed that between the two 2018 cohorts, Chinese L1 students were far more likely to use the personal pronouns *we*, *I*, and *us* in their article critiques. In addition, three of the five most frequently used parts of speech by the Chinese L1 students were personal pronouns, proper nouns, and adverbs of location. The table above and the discussion of stance suggest that vocabulary selection and use in writing by Chinese L1 beginner university students may trail their international peers in the BIBA Freshman English class.

## 4.2. Syntactic Complexity

To examine syntactic complexity, we used both the L2 Syntactic Complexity Analyzer (Ai & Lu, 2013) and the tabular assessments of parts of speech established by the CLAWS tagger in WMatrix4®. Our findings echoed these previous studies when we used Ai and Lu's (2013) Web-based *L2 Syntactic Complexity Analyzer* (L2SCA; <https://aihaiyang.com/software/l2sca/>) to examine randomly chosen 1000-word samples of BIBA freshman critiques for 2017 and 2018 with reviews from more advanced student writers in MICUSP. L2SCA has fourteen indices that predominantly look at clauses and T-units and verb and coordinate phrases and complex nominals. In Table 3, we omitted findings for T-units, following Biber et al. (2011).

**Table 3.** Syntactic Complexity Comparisons Using L2SCA on Thousand-word Samples

	S	VP	CP	CN	MLS	MLC	CP/C	CN/C
2017 1000-word Sample	46	132	15	113	20.63	9.40	0.149	1.12
2018 1000-word Sample	56	163	26	112	17.82	8.91	0.232	1.00
2018 Chinese 1000-word Sample	42	155	19	116	22.88	8.74	0.173	1.05
2018 Non-Chinese 1000-word Sample	46	132	34	130	21.98	10.87	0.366	1.40
MICUSP 1000- word Sample	30	105	17	117	33.23	13.47	0.230	1.58

*Legend:* S Sentence; VP Verb phrase; CP coordinate phrase; CN complex nominal; MLS Mean length of sentence; MLC mean length of clause; CP/C coordinate phrase per clause; CN/C complex nominal per clause

We see these figures as suggestive of trends; slightly different totals will be obtained with other samples from the same small corpora, but they signal similar differentiation. The number of sentences

is of immediate pedagogical interest. It suggests that advanced, predominantly English L1 graduate students whose papers were collected for MICUSP, had fewer but much longer sentences typically containing fewer verb phrases. While at the same time incorporating more elaborated clauses with a much greater number of complex nominals per clause. In addition, the table also shows that non-Chinese students demonstrate higher complex structure in their writing samples than the Chinese L1 students. Interestingly, the non-Chinese first-year students, who were typically first-language English, elaborated their clauses with coordinate phrases, e.g., *most of them are useful and easy to do*, as well as complex nominals. The finding of syntactic complexity among the five writing sample groups is in line with Ai and Lu's (2013) report that EALs are more likely to have shorter mean lengths of clauses and sentence complexity (Table 3).

Learning to read, interpret, develop, and incorporate complex nominals as both post- and especially pre-modifiers is, as Ruan (2016) demonstrates, essential for second language majors in business and business-related fields. Using a corpus of annual reports and CEO letters from 16 banks, Ruan emphasizes the importance of structures such as "*a very successful general insurance business*" as an example of premodification, with prepositional phrases as "the majority of postmodifiers in various registers" (Ruan, 2016, pp. 75-76). Of particular interest is the correlation of elements of premodification with semantic and pragmatic associations (p. 85), which leads Ruan to emphasize the pedagogical importance of furnishing second-language learners with business-related readings and corpora of authentic materials.

### 4.3. Prepositions

In the development of post-modification elaboration, different prepositions and prepositional phrases are frequently used, but prepositions cause many issues for EALs. Nacey (2013) comments that since prepositions are polysemous, it can be challenging for EAL to distinguish one from another, mainly since they can function in multiple word classes and are frequently metaphorical. We reviewed single-word and phrasal, or complex prepositions, using WMatrix4® on the full text of each cohort. Even without counting *of* and *with*, prepositions were in the five most frequent parts of speech for all three corpora of critiques. Table 4 compares the ten most frequently used parts of speech in final papers by non-Chinese and Chinese freshman BIBA students, suggesting that Chinese EAL writing was slightly more clausal, using conversational parts of speech (POS) such as lexical and modal verbs more than non-Chinese writing. Similar results were also reported in Bychkovska and Lee's (2017) study. In the table below, boldface identifies differences in the writing cohort.

**Table 4.** Ten Most Frequent Parts of Speech Used by L1 and EAL Writers in BIBA First-year English Class

Variable	Not Chinese: More Phrasal	Chinese only: More Clausal
Total words	25976	16329 (62.9% of all words)
Top 10 POS	NN1, JJ, NN2, II, AT, VVI, CC, RR, <b>AT1, TO</b> Singular common noun General adjective Plural common noun Prepositions Article the Infinitive verb And/or Coordinating Conjunction Adverb <b>Singular article a/an</b> <b>To (Infinitive and/or directional marker)</b>	NN1, JJ, NN2, AT, II, VVI, RR, CC, <b>VV0, VM</b> Singular common noun General adjective Plural common noun Article the Prepositions Infinitive verb Adverb And/or Coordinating Conjunction <b>Lexical Verb</b> <b>Modal Verb</b>



For all cohorts, 2017, 2018 (Combined, Chinese-only, and non-Chinese) and MICUSP critiques, *about, as, by, from, in, on,* and non-infinitival *to* comprised 75% or more of all single-word prepositions. What Miller (2020) terms as idioms (see her Appendix B, Miller, 2020, p. 12) were similar as well: those used four or more times in each cohort's corpus included *above all, at the same time, in the end, in other words, in my opinion, in the past few years* and *on the other hand*. In context, each was typically used adverbially, either prefacing a clause or a complete sentence. The wording of each could be slightly changed (*in my honest opinion, in my considered opinion, in my earlier opinion,* and so on). Complex prepositions, on the other hand, cannot incorporate insertions or changes. Complex prepositions used four or more times by each cohort include: *according to, because of, due to, instead of, rather than,* and by one writer in the not-Chinese cohort, *thanks to*. Most frequently, these complex prepositions were related to some aspect of citation and the integration of sources.

#### 4.4. Stance

As we have noted earlier, hedging – focused here on modality - is extensively studied in research on stance in academic or university writing, as exemplified. As Nesi and Gardner (2018) state, epistemic modality is often a key to persuasion. More recently, pragmatics constructs have been studied, specifically related to politeness: for example, Qian and Pan (2019) reported on the pragmatic use of modals in business letters written by Chinese learners. They based their analysis on a corpus of 600 business letters they collected from university students in Hong Kong and Shanghai: English is widely used in Hong Kong where it is an official language; in Shanghai, it is foreign, if frequently used language for business communication (2019, pp. 21-22). Since letter-writing in business is most commonly designed to be persuasive in nature, examining modals within phraseological contexts allows the researchers to compare each group's use of individual modals and investigate distribution patterns. When we reviewed the use of modals in each of our cohorts (using WMatrix4® as did Qian & Pan, 2019), we identified similar usage for individual modals: *can, will,* and *would*. These were the most frequent and were not always diplomatically polite (as in 'You *will* do X'). In addition, the individual modal *must* is the least favorite among all groups in our study. The similar results of the usages of modals identified in the present study and Qian and Pan's (2019) might be attributed to the shared linguistic and cultural background.

**Table 5.** Modals across Cohorts

Modal	Typical Category	BIBA 2017	BIBA 2018 Chinese only	BIBA 2018 Non-Chinese	Critiques MICUSP
Can	Epistemic/ability	0.77	1.15	1.06	0.48
Could	Epistemic/ability	0.10	0.07	0.07	0.13
May	Epistemic/ability	0.13	0.15	0.06	0.15
Might	Epistemic/ability	0.10	0.19	0.06	0.09
Will	Volition/prediction	0.59	0.51	0.63	0.26
Would	Volition/prediction	0.08	0.10	0.08	0.30
Must	Deontic/obligation	0.08	0.04	0.07	0.10
Should	Deontic/obligation	0.26	0.31	0.11	0.10
	Number of modal types	10	13	8	14
	Frequency of modal types	2.12	2.57	2.16	1.66

Hedges are, of course, composed of more than modal verbs or phrases. All four cohorts used the following, as identified by *TextInspector*: *almost, could, likely, may, might, often, possible, relatively, usually,* and *would*. In addition, students in both 2018 first-year cohorts added *approximately* but varied in their use of -ly adverbs suggesting degrees of certainty, possibility, and frequency; the

advanced writers in MICUSP expanded the range of tokens in the same category and added verbs like *estimate*.

## 5. Conclusions: Implications for Future Pedagogical Development

In this study, we employed a corpus-based approach to explore differences in the language use of Chinese and non-Chinese university students' critiques written for a freshman English course in the BIBA program. Further, we examined how these student writers compare with high-rated L1 students from the MICUSP corpus. Through the use of various online analyses, we found that by the end of the first year, the students, most of whose entrance scores had placed them in CEFR B1, had increased their vocabulary to the B2 or C1 levels. In addition, they now indicated that they possessed an ability to identify relevant sources. However, after examining the grammatical features and interactional metadiscourse used by the student writers, we discovered differences between Chinese and non-Chinese students in their language use. For instance, Chinese student writers had less variety of content words. That is, they tended to repeat a more restricted set of words, illustrated by their dependence on only a few prepositions. In addition, critiques from the Chinese writers contained shorter clauses and sentences, and the mean value of their complex nominal and coordinate phrases was smaller than non-Chinese and MICUSP writers. Finally, in terms of lexical bundles, Chinese student writers showed preferences for clausal bundles instead of phrasal bundles, which are often found in texts of skilled writers in previous literature. Findings from this study highlight the importance for second language teachers to be aware of the gap between Chinese and non-Chinese students regarding their lexical richness, syntactical complexity, and facility with interactional metadiscourse.

The study is not without limitations. The learner corpus size selected for this study is small, so that future studies can benefit from collecting more student writing samples from students with different L1 backgrounds. In addition, it will be helpful to examine the results of various written topics, tasks, or learning contexts in EAP or more ESP fields. Despite the limitations, the current study's findings add to understanding the differences in language use between native English speakers and learners with different L1 backgrounds.

The findings of the current study offer insightful information for future pedagogical decisions. The lexical gap between student groups calls for the implication of a relevant curriculum to enhance students' linguistic knowledge. For instance, for a Business English course, we should continue exploring ways to increase single- and multi-word vocabulary items across business areas to incorporate finance, economics, and marketing technologies. We can also draw on additional learner or professional corpora as resources for our course design to encourage students to employ more complex syntactic structures. In addition to the implicit approach of providing writing samples from successful writers in both native and non-native English, these samples support writing instruction that can benefit students with diverse learning needs. Designing lessons that focus on grammatical forms and content and illustrating phrasal bundles and phrase-based complex items will further be enhanced with the samples from different corpora. Moreover, the learners' writing samples can serve as an example corpus for language teachers to track and better understand students' writing development and provide information for designing courses and materials that cater to specific language needs in different professions.

## References

- Ai, H., & Lu, H. (2013). A corpus-based comparison of syntactic complexity in NNS and NS university students' writing. In Diaz-Negrillo, A., Baillier, B., & Thompson, P. (Eds.), *Automatic Treatment and Analysis of Learner Corpus Data* (pp. 249-264). Philadelphia: John Benjamins.
- Aull, L. (2019). Linguistic markers of stance and genre in upper-level student writing. *Written Communication, 36*, 267-295.
- Bargiela-Chiappini, F., & Zhang, Z. (2013). Business English. In Paltridge, B., & Starfield, S. (Eds.), *The Handbook of English for Specific Purposes* (pp. 193-211). West Sussex: John Wiley & Sons.
- Bax, S., Nataksuhara, F., & Waller, D. (2019). Researching L2 writers' use of metadiscourse markers at intermediate and advanced levels. *System, 83*, 79-85.
- Biber, D., Anthony, M., & Gladkov, K. (2007). Rhetorical appeals in fundraising. In Biber, D., Connor, U., & Upton, T. (Eds.), *Discourse on the Move: Using Corpus Analysis to Describe Discourse Structure* (pp. 121-151). Philadelphia: John Benjamins.
- Biber, D. (1988). *Variation Across Speech and Writing*. Cambridge: Cambridge University Press.
- Biber, D., Gray, B., & Poonpon, K. (2013). Pay attention to the phrasal structures: going beyond T-units—a response to WeiWei Yang. *TESOL Quarterly, 47* (1), 192–201.
- Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly, 45*(1), 5-35.
- Bychkovska, T., & Lee, J. (2017). At the same time: Lexical bundles in L1 and L2 university student argumentative writing. *Journal of English for Academic Purposes, 30*, 38-52.
- Carlsen, C. (2018). The adequacy of the B2 Level as university entrance requirement. *Language Assessment Quarterly, 15*, 75-89.
- Chan, C. (2019). Long-term workplace communication needs of business professionals: Stories from Hong Kong senior executives and their implications for ESP and higher education. *English for Specific Purposes, 56*, 68–83.
- Chan, C. (2018). Proposing and illustrating a research-informed approach to curriculum development for specific topics in business. *English for Specific Purposes, 52*, 27–46.
- Cheng, W., Lam, P., & Kong, K. (2018). Learning English through workplace communication: Linguistic devices for interpersonal meaning in textbooks in Hong Kong. *English for Specific Purposes, 55*, 28-39.
- Gardner, S., Nesi, H., & Biber, D. (2019). Discipline, level, genre: Integrating situational perspectives in a new MD analysis of university student writing. *Applied Linguistics, 40*, 646-674.
- Gherdan, M. (2019). Hedging in academic discourse. *Romanian Journal of English Studies, 16*, 123-126.
- Grabe, W., & Zhang, C. (2016). Focus on texts and readers: Linguistic and rhetorical features. In Manchén, R., & Matsuda, P. (Eds.), *Handbook of Second and Foreign Language Writing* (pp. 245-266). Berlin: De Gruyter.
- Ho, V., & Li, C. (2018). The use of metadiscourse and persuasion: An analysis of first year university students' timed argumentative essays. *Journal of English for Academic Purposes, 33*, 53-68.
- Jitpraneechai, N. (2019). Noun phrase complexity in academic writing: A comparison of argumentative English essays written by Thai and native English university students. *Language Education and Acquisition Research Network Journal, 12*, 71-88.
- Kim, J., & Nam, H. (2019). How do textual features of L2 argumentative essays differ across proficiency levels? A multidimensional cross-sectional study. *Reading and Writing, 32*, 2251-2279.
- Kirkpatrick, A., & Xu, Z. (2012). *Chinese Rhetoric and Writing: An Introduction for Language Teachers*. Fort Collins, CO: The WAC Clearing House.
- Lee, J., & Deakin, L. (2016). Interactions in L1 and L2 undergraduate student writing: Interactional

- metadiscourse in successful and less-successful argumentative essays. *Journal of Second Language Writing*, 33, 21-34.
- Lu, X. (2012). The relationship of lexical richness to the quality of ESL Learners' oral narratives. *The Modern Language Journal*, 96(2), 190-208.
- Miller, J. (2020). The bottom line: Are idioms used in English academic speech and writing? *Journal of English for Academic Purposes*, 43, 1-14.
- Nacey, S. (2013). *Metaphors in Learner English*. Philadelphia: John Benjamins.
- Nesi, H., & Gardner, S. (2018). The BAWE corpus and genre families classification of assessed student writing. *Assessing Writing*, 38, 51-55.
- Nesselhauf, N. (2004). Learner corpora and their potential for language teaching. *How to Use Corpora in Language Teaching*, 12, 125-156.
- Qian, D., & Pan, M. (2019). Politeness in business communication: Investigating English modal sequences in Chinese Learners' letter writing. *RELC Journal*, 50, 20-36.
- Ruan, Z. (2016). Lexical bundles in Chinese undergraduate academic writing at an English medium university. *RELC Journal*, 48, 327-340.
- Schieber, D., & Robles, V. (2019). Using reflections to gauge audience awareness in business and professional communication courses. *Business and Professional Communication Quarterly*, 82, 297-316.
- Staples, S., Egbert, J., Biber, D., & Gray, B. (2016). Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline, and genre. *Written Communication*, 33(2), 149-183.

## THE AUTHORS

Yu-Chi Wang is an assistant professor in the Foreign Language Center at Feng Chia University in Taiwan. She received her doctoral degree in Foreign Language and ESL Education program at the University of Iowa and a master's degree in Reading Education at the University of Pittsburgh. Her research interests include learner identity issues and topics related to teaching methodology and second language acquisition in the TESOL field.

Richard Hill Davis is an associate professor in the Foreign Language Center at Feng Chia University in Taiwan. He received his doctoral degree in English Language at Glasgow and a master's degree in TESOL at National Kaohsiung Normal University. His research interests include ESP, EMI, Academic Writing and issues related to teaching methodology and second language acquisition in the TESOL field.

## THE AUTHORS' ADDRESSES

### First and Corresponding Author

#### Yu-Chi Wang

Assistant Professor  
 Foreign Language Center  
 Feng Chia University  
 No. 100, Wenhua Rd. Xitun Dist., Taichung City, TAIWAN  
 Email: yuchiwang77@gmail.com

### Co-author

#### Richard Hill Davis

Associate Professor  
 Foreign Language Center  
 Feng Chia University  
 No. 100, Wenhua Rd. Xitun Dist., Taichung City, TAIWAN  
 Email: rhdavis@fcu.edu.tw

Received: 14 June 2021

Received in Revised Form: 19 November 2021

Accepted: 6 December 2021