A Contrastive Interlanguage Analysis of Lexical Bundles in English as a Foreign Language Writing: L1 Chinese, Japanese, and Korean

Randy Appel
Waseda University

Lewis Murray
Kanazawa University

Correct and register-appropriate use of frequently recurrent word sequences (e.g., lexical bundles) plays an important role in proficient linguistic output. However, second language (L2) writers' use of these multiword items is still insufficiently understood, particularly in relation to the influence of first language (L1) background. This exploratory study analyzed a learner corpus of 420 argumentative essays to determine how lexical bundles were used by L2 English academic writers from 3 L1 backgrounds (Chinese, Korean, Japanese) to identify intra-group tendencies and intergroup production differences. A contrastive interlanguage approach identified unique tendencies related to functional categories and individual lexical bundles for each L1 group. Findings include relative overuse of text-oriented bundles by L1 Chinese writers, overuse of participant-oriented bundles by L1 Japanese writers, and a general tendency to underuse of lexical bundle types and tokens by L1 Korean writers of L2 English. Methodological and pedagogical implications of these findings are discussed.

Lexical bundles(単語連鎖)のように高頻度で使用される語の連続を、適切に、正しいレジスターで用いることは、熟練した言語の産出に不可欠である。しかし、第二言語学習者によるその使用の実態は、特に母語の影響との関連では十分に理解されていない。本研究では、420の学術的文章からなるコーパスの分析を通して、母語背景(中国語、韓国語、日本語)をもつ英語の第二言語学習者によるlexical bundlesの使用を調べ、母語集団内の傾向や集団間でのその産出の違いを明らかにする。対照中間言語分析によって、談話機能や個別のlexical bundlesに関して、

https://doi.org/10.37546/JALTJJ45.1-2

JALT Journal, Vol. 45, No. 1, May 2023

各母語集団独自の傾向があることがわかった。研究の成果として、中国語母語の書き手がテキスト志向のbundlesを、日本語母語の書き手が参与者志向のbundlesを、それぞれ比較的多用する傾向があることや、一般に韓国語母語の書き手にはlexical bundlesのタイプやトークンの使用頻度が低い傾向がみられることを示す。これらの方法論的・教授法的含意についても論じる。

Keywords: formulaic sequences; L1 differences; multiword structures; second language writing

Research into the use of Formulaic Sequences (FSs), defined as multiword structures believed to be stored and produced as single units (Wray, 2002) has grown exponentially over the past several decades. These studies cover a range of foci, yet one of the main findings is that FSs play a vital role in fluent and proficient language use (e.g., Chen, 2019; Wray, 2002). Despite their importance, contrastive research into how these structures are used by second language (L2) English writers of differing first language (L1) backgrounds is limited. In particular, studies examining use of lexical bundles (LBs), a frequency-based approach to the identification of FSs, by writers of varying L1 backgrounds are rare. Furthermore, studies that have examined LBs in this way have often contained methodological issues preventing distinctions between L1 specific and universal tendencies.

As a result, further studies examining more closely controlled and comparable corpora are needed to better understand how LBs are employed by L2 English writers of varying L1 backgrounds. Findings may lead to improved identification and distinction of L1-related and universal production tendencies that could be used to inform pedagogic interventions aimed at English as a Foreign Language (EFL) and English as a Second Language (ESL) users. The current exploratory study therefore analyzed 420 EFL essays by writers of three distinct L1 backgrounds: Chinese, Japanese, and Korean, with English proficiency, writing conditions, and topic controlled across groups.

Literature Review

Formulaic Sequences & Lexical Bundles

FSs are increasingly recognized as a crucial aspect of fluent and proficient language use, due in large part to the widespread use of corpus informed research that has helped drive growth in this area. Thus far, scholars have shown that FSs are prevalent in L1 speech and writing (e.g., Schmitt, 2004), aid in perceptions of fluency (e.g., Wray, 2002), and ease processing and production burdens associated with unplanned linguistic output (e.g., Kuiper, 1996).

Whereas definitions of FSs can vary depending on the goals of the researcher, LBs carry a more stable definition that results in greater interstudy comparability. Introduced by Biber et al. (1999), LBs are simply defined as multi-word strings (often four words in length) that meet minimum frequency and range criteria. This quantitative focus means LBs often cross semantic and syntactic boundaries and may not hold the same psycholinguistic status as wholly stored and produced FSs. However, like FSs, LBs contribute to perceptions of linguistic proficiency (e.g., Shin, 2019) and distinguish L1 from L2 users (e.g., Lu & Deng, 2019).

The LB approach has grown from a relatively niche method to one widely used to analyze L1 and L2 discourse across a range of genres and registers. Examples include Hyland (2008), who used a 3.5-million-word corpus of academic texts to reveal production tendencies that help distinguish scholarly disciplines (e.g., engineering, biology, business, applied linguistics). Similarly, Durrant (2017) analyzed the British Academic Written English (BAWE) corpus identifying distinct LB production patterns that differentiate hard and soft sciences.

Lexical Bundles in L2 English Writing

As with many forms of corpus-driven/based research, early LB studies commonly focused on L1 English discourse. However, this focus has gradually shifted to examine structures by L2 English writers of varying proficiencies (e.g., Appel & Wood, 2016; Chen, 2019). In general, these studies aim to identify production patterns that distinguish high- and low-level writing, with findings informing teaching interventions aimed at improving proficiency. For example, Chen (2019) used a large collection of essays from the International Corpus Network of Asian Learners of English (ICNALE) to reveal that EFL writers at higher proficiency levels used a wider range of LBs in their texts. Similarly, Appel and Wood (2016) examined data from a frequently used standardized English proficiency test to highlight how lower-level writers favor LBs indicating personal stance.

Although previous studies generally grouped L2 English learners from varying L1 backgrounds together in hopes of identifying more widely applicable findings, the identification of production patterns that distinguish L2 users on the basis of their L1 has been growing in popularity. This area of LB research follows a more general trend in corpus informed studies of ESL/EFL writing that aims to identify L1 specific and universal production tendencies in L2 output (e.g., Jarvis & Pavlenko, 2008), often focusing on L1 Chinese learners of L2 English (e.g., Bychkovska & Lee, 2017; Chen &

Baker, 2010). For example, Bychkovska and Lee (2017) compared post-secondary English texts produced by L1 Chinese and L1 English writers to reveal that L1 Chinese students made use of more LB types and tokens than L1 English writers. These findings were attributed to a higher number of conversational LBs in L1 Chinese writing and heavy dependence on direct translation equivalents.

Although a focus on L1 Chinese EFL writing continues, additional L2 English users have also been investigated. Allen (2011), for instance, used a corpus of EFL writing to show that L1 Japanese writers of English tended to overuse LBs that had translation equivalents in their native language. For example, with reference to stance bundles, the author notes the high frequency of *it can be said*. Comparing L1 Korean EFL writers and native English users, Shin (2019) discovered that L1 Korean students displayed a greater tendency for stance and discourse organizing LBs.

Limitations in Previous Research

The above-mentioned studies add valuable knowledge regarding how LBs are used by various populations of L2 English writers, yet several inherent limitations persist. Notably, most of this research has used one-to-one contrasts involving a single group of L2 English writers in comparison to a reference corpus of L1 English (e.g., Allen, 2011; Chen & Baker, 2010; Shin, 2019). As this approach does not include additional L1 groups for comparison purposes, conclusions regarding whether identified production patterns are L1-related, or common to all L2 English writers, are impossible. Furthermore, L1/L2 comparisons often involve target language proficiency differences that may result in the misattribution of findings. Thus, moving away from L1/L2 comparisons in favor of contrasts targeting the interlanguage of multiple L1 groups may be more valuable (Ortega, 2011).

Several studies have begun to involve multiple populations of L2 English writers in their research (e.g., Appel & Murray, 2020; Karabacak & Qin, 2013; Paquot, 2017). Unfortunately, these studies have often failed to adequately control for proficiency and writing conditions and/or used extremely small sample sizes. For instance, in Karabacak and Qin (2013), only 17 samples from each of the L1 groups (Turkish, Chinese, and English) were used. Thus, it is difficult to make generalizable statements that could apply more broadly to each population of writers. Although Paquot (2017) analyzed a much larger collection of writing and used an innovative approach to highlight potential L1 related production tendencies, reliance on the International Corpus of Learner English (ICLE) could be

seen as a limitation, as the writing comes from post-secondary institutions with varying academic standards, writing conditions, and target language proficiencies. An examination of a small collection of writing from the ICLE found a range of B2 to C2 on the Common European Framework of Reference for languages (Granger & Thewissen, 2005) suggesting results from studies using the ICLE should be taken with caution. Similarly, the three corpora of ESL writing analyzed by Appel and Murray (2020) were comprised only of 'passing grade' papers, but the authors acknowledged that this 'pass/fail' distinction may have been overly broad in terms of controlling for proficiency, thus negatively impacting findings.

The Current Study

With limitations of previous research in mind, the present study aimed to use a more closely comparable collection of L2 English samples to perform a contrastive interlanguage analysis of L1 Chinese, Japanese, and Korean EFL writing. These L1s were chosen as they represent three of the most common groups of L2 English users from East Asian countries studying in English medium universities. Thus, a better understanding of these students' writing could offer benefits in terms of more targeted instruction that better addresses the needs of each group. The main research question providing focus to this study is provided below:

RQ. How do L1 Chinese, Japanese, and Korean EFL students make use of LBs in their academic English writing?

As an exploratory study, we focused on identification of potential L1 related production tendencies through the analysis of LBs. Although the discussion proposes factors that may explain our findings, it is beyond the scope of this paper to more definitively identify specific root causes. It is hoped that future research will build on the present study by incorporating analyses of L1 corpora, translation equivalents, pedagogic materials, pedagogic approaches, and L1 congruence to better understand the role each of these factors may play in the highlighted results.

Method

Corpora

Data were assembled from version 2.3 of the Written Essay Module of the ICNALE. The ICNALE is composed of essays and speeches by post-secondary students from 10 countries using standardized data collection procedures

that include common topics, writing conditions, access to materials, and allotted time. This corpus was specifically designed to facilitate studies focused on contrastive interlanguage analyses. In total, 5,600 essays from 2,800 writers are included in the ICNALE; however, there is substantial variance in terms of number of samples and assessed proficiency among each of the L1 sub-corpora. For example, although the Japanese and Chinese sections of ICNALE both contain 400 samples, only 50 of these have been assessed to be A2 in the L1 Chinese section and 154 of these are assessed at this level in the Japanese section.

For the current study, 140 essays from each of the three previously mentioned L1 groups (Chinese, Japanese, Korean) were gathered from the B1 level[1], as this proficiency band contained a relatively large pool of data from which to draw. All essays were written to address the same writing prompt: *It is important for college students to have a part-time job* (agree or disagree) as a way of controlling for any potential topic influence. As can be seen in Table 1, Chinese writers produced the longest average essays and showed the greatest standard deviation, yet all three groups were comparable in terms of total corpus size and mean essay length.

Table 1

Descriptive Statistics of the Corpora

	Chinese (n = 140)	Japanese (n = 140)	Korean (n = 140)
Total running words	34,575	31,892	31,988
Mean (range)	245 (195-338)	226 (176-302)	227 (189-326)
Standard deviation	34	23	27

Extraction Criteria

Range & Frequency

Range and frequency are the main identifying criteria for LBs, yet values vary from study to study. Minimum range is a means of avoiding idiosyncratic language use from a minority of texts/users which could skew results by misrepresenting general tendencies. Previous studies have often set range as either a raw number (e.g., Chen & Baker, 2010; Shin, 2019) or percentage (e.g., Appel & Murray, 2020; Hyland, 2008) of the total number of texts. As raw numbers can be influenced by the total number of samples in each

corpus (i.e., achieving a 5-text minimum in a corpus of 50 essays may be more difficult than achieving this same number in a corpus of 500 essays), a percentage threshold for range was used in the current study. As the main function of range is the elimination of idiosyncratic tendencies, 10% (14 texts from each corpus) was used to achieve this goal.

In instances where large corpora or corpora of substantially different sizes are analyzed, a normalized frequency (typically a value per million words) is used. Conversely, in studies examining smaller corpora, or those with more comparable word counts, a raw frequency is more commonly applied. Because the three corpora in the current study are of a comparable size, and all essays were relatively short (approximately 230 words, on average), the previously established minimum range figure of 14 was also applied as the frequency criterion. Thus, any bundle appearing in at least 10% of texts from any L1 group (14 occurrences) would fulfil both the range and frequency criteria.

Length

For sequence length, 4-word bundles are common as this often produces a manageable set of items for analysis (e.g., Chen & Baker, 2010), shorter sequences are contained within their boundaries (Cortes, 2004) and they offer relatively clear functional roles (Hyland, 2008). However, an exclusive focus on 4-word structures has been criticized in recent years as leading to potential misidentification because longer and shorter sequences are hidden from analysis (e.g., Adel & Erman, 2012; Appel & Trofimovich, 2017).

In the current study, we began by extracting all 3-word sequences, with instances of contracted forms treated as two words (e.g., *don't*, *won't*). However, substantial overlap suggested the presence of longer repeated structures. Therefore, target length was expanded to include any sequence meeting the aforementioned frequency and range, regardless of length. In doing so, it was possible to identify highly frequent LBs of up to 14 words in length. To eliminate the presence of partially overlapping sequences, all extracted items were reviewed, and shorter structures embedded in longer sequences were eliminated before beginning the analyses. For example, *a part time* was identified in the L1 Chinese corpus as a 3-word LB. The bundle *a part time job* was also identified, superseding the 3-word bundle, which in turn was superseded by the 5-word bundle *a part time job in*. As this 5-word bundle met the previously established extraction criteria, this 5-word sequence was retained but the shorter LBs it contained (*a part time*, *a part time job*) were eliminated from subsequent analyses.

Prompt-Related Bundles

Contrary to many earlier studies that used writing from a range of topics and/or genres, the current study examined argumentative essays addressing one common prompt. Thus, although previous scholars have often removed prompt/topic related LBs from their analyses, this was unnecessary in the present research. However, although bundles containing the topic-related words 'part-time', 'college' and 'students' were included throughout the analysis, they are not included in the analysis of individual items in Tables 4, 5, and 6 as a way of providing focus and allowing greater emphasis on the discussion of more interpretable findings.

Analysis

Extraction and analysis of LBs followed three steps. First, LBs meeting the identification criterion were extracted from each corpus. These were then reviewed to eliminate overlapping sequences (i.e., shorter sequences that were constituent in longer LBs). Second, cleaned lists of 3-14-word sequences were functionally classified using Hyland's (2008) categorization This taxonomy was used as it builds on Biber et al.'s (2004) classifications, has been argued to be better suited to academic texts, and follows our previous research (Appel & Murray, 2020) which can lead to greater comparability. This classification system includes three major functional categories. Research-oriented bundles aid the explanation of real-world occurrences, often through direct reference to concrete objects and abstract concepts (e.g., ***The part-time job that*** *most students have*); text-oriented bundles aid in the organization of discourse by helping to guide the reader, often with signposting language (e.g., ***First of all****, …*); and, participant-oriented bundles are writer/reader-focused and often serve to provide statements that make clear the writer's personal opinions (e.g., ***I think that*** *this should be stopped*). Each major functional category also contains distinct subcategories (see Hyland, 2008). Following functional classifications, loglikelihood statistics were used to highlight significant intergroup production differences related to functional (sub)categories and individual LBs. In all cases, findings are only reported as significant if these contrasts yielded $p < .01$.

We begin by providing a brief overview of general findings regarding frequency of various sequence lengths in the three corpora. This is followed by analysis of the most commonly used LBs by each L1 group. Finally, functional category and individual item comparisons are used to highlight significant intergroup production differences.

Results

General Findings

L1 Japanese were the most frequent users of LBs as a whole, with a total of 51 types and 1,138 tokens (Table 2). L1 Chinese writers closely followed L1 Japanese writers in terms of total LB occurrences and L1 Korean writers were the least frequent users. Most striking is the high number of longer LBs (greater than 5-words) in the L1 Japanese corpus, which suggests significant overlap and a high level of intra-group similarity. In contrast, L1 Korean writers seem to possess the greatest intra-group variance as both type and token counts across nearly all sequence lengths were comparatively low.

Table 2
3-word to 14-word Bundles by L1 Group

Bundle Length	Chinese	Japanese	Korean
3-word	25 (566)	29 (665)	20 (465)
4-word	11 (236)	11 (238)	11 (274)
5-word	8 (215)	1 (33)	4 (111)
6-word	3 (66)	5 (92)	1 (21)
7-word	-	2 (32)	-
8-word	-	1 (15)	-
12-word	1 (24)	-	1 (23)
13-word	-	1 (44)	-
14-word	-	1 (19)	-
Total	48 (1,107)	51 (1,138)	37 (894)

Note. Type counts are listed outside of parentheses with token counts listed within parentheses. All 9-, 10- and 11-word bundles were constituent in longer bundles.

Functional Analysis

To ensure consistency in functional assignment (Table 3), classifications were performed independently by the two authors of this study before reconvening to discuss discrepancies. Interrater reliability for these initial classifications was 91% (agreement on 124 of 136 total LBs). Full agreement was achieved through joint discussion.

In the following analyses, the terms *overuse* and *underuse* are used to refer to instances of statistically distinct production by one L1 group in comparison to the other two L1s. These terms are used in a relative manner and should not be seen as an indication of improper use. Where pedagogic implications are given, this is explicitly stated.

Table 3
Lexical Bundles by Functional Category

	Chinese	Japanese	Korean
Research-oriented	850 (77%)	697 (61%)	778 (87%)
Location	55 (5%)	27 (2%)	33 (4%)
Procedure	276 (25%)	132 (12%)**	222 (25%)
Quantification	138 (12%)**	66 (6%)	48 (5%)
Description	266 (24%)	245 (22%)	189 (21%)**
Topic	115 (10%)**	227 (20%)	286 (32%)
Text-oriented	102 (9%)**	52 (5%)**	16 (2%)**
Transition	102 (9%)**	52 (5%)**	16 (2%)**
Participant-oriented	155 (14%)**	389 (34%)**	100 (11%)**
Stance	155 (14%)**	389 (34%)**	100 (11%)**
Total	1,107 (100%)	1,138 (100%)	894 (100%)**

Note. $^{**}p < .01$; Only sub-categories with LB occurrences are listed. Type counts are listed outside of parentheses with percent of token counts listed within parentheses.

Two major functional categories (*text-oriented*, *participant-oriented*) displayed significant L1 related production differences (Table 3). Furthermore, loglikelihood statistics indicated that each L1 held a unique tendency related to the *text-oriented* category, with L1 Chinese writers the most frequent users. The *participant-oriented* category also displayed similar results, with each L1 making use of these items in a statistically unique manner. However, in this case, L1 Japanese writers were the most frequent users.

Within the *research-oriented* category, the *procedure*, *quantification*, *description*, and *topic* subcategories showed significant intergroup production differences. For the *procedure* subcategory, L1 Japanese writers were the least frequent users; L1 Chinese overused *quantification* yet underused

the *topic* subcategory; *description* was underused by L1 Korean writers. In terms of total use across all categories, L1 Korean writers were found to underuse LBs (tokens).

Individual Lexical Bundle Analysis

L1 Chinese

L1 Chinese EFL writing contained the highest number of unique L1-related production tendencies (19) for individual LBs. These items were roughly split between overused (10) and underused (9) items. In relation to significant functional category differences identified in Table 3, only one item (*all kinds of*) could be identified as contributing to the relative overuse of the quantification subcategory (i.e., LBs used to describe amounts). However, 4 LBs from the text-oriented category (*all in all*, *at the same time*, *last but not*, *what's more*) helped to explain the relative overuse of this particular functional category.

Table 4
Individual Item Overuse/Underuse by L1 Chinese Writers

Category	Bundle	Chinese	Japanese	Korean
Research-oriented: Location	*in the society*	17	4	2
Research-oriented: Purpose	*be able to*	1	21	16
Research-oriented: Purpose	*for us to*	34	12	0
Research-oriented: Purpose	*I want to*	4	28	35
Research-oriented: Purpose	*they (have/ want) to*	2/4	15/14	19/17
Research-oriented: Quantification	*a lot of (money)*	51 (2)	116 (27)	77 (19)
Research-oriented: Quantification	*all kinds of*	14	1	1
Research-oriented: Description	*importance of money*	0	17	8
Text-oriented: Transition	*all in all*	15	0	0
Text-oriented: Transition	*at the same time*	17	3	2
Text-oriented: Transition	*but it is*	3	13	16

Category	Bundle	Chinese	Japanese	Korean
Text-oriented: Transition	*last but not*	16	0	0
Text-oriented: Transition	*what's more*	24	0	0
Participant-oriented: Stance	*as far as I*	15	0	1
Participant-oriented: Stance	*I think that*	8	96	30
Participant-oriented: Stance	*in my opinion*	34	4	8
Participant-oriented: Stance	*we all know*	15	0	0

L1 Japanese

Significant differences for individual LBs among L1 Japanese writers covered all three major functional categories, yet were primarily related to the *participant-stance* subcategory. With all items from this category indicating relative overuse, findings in Table 5 help to explain the previously identified functional overuse by L1 Japanese EFL writers. Overused items are primarily used to indicate writers' personal opinions, with several making explicit mention of the writer by way of 'I'. In terms of relative underuse of the *research-procedure* subcategory (used to detail processes) highlighted above, two LBs (*for us to*, *to get a*) help to explain this.

Table 5

Individual Item Overuse/Underuse by L1 Japanese Writers

Category	Bundle	Chinese	Japanese	Korean
Research-oriented: Purpose	*for us to*	34	12	0
Research-oriented: Purpose	*to get a*	12	2	22
Research-oriented: Quantification	*a lot of (things)*	51 (1)	116 (20)	77 (3)
Research-oriented: Description	*we can't*	10	25	2
Research- oriented: Topic	*the statement that*	1	17	0
Research- oriented: Topic	*with this statement*	1	14	0
Text-oriented: Transition	*and so on*	13	52	13

Category	Bundle	Chinese	Japanese	Korean
Participant-oriented: Stance	*I agree with the statement*	1	33	0
Participant-oriented: Stance	*I agree with this*	1	25	1
Participant-oriented: Stance	*I think it is*	19	53	15
Participant-oriented: Stance	*I think that (it is important for)*	8 (2)	96 (16)	30 (3)
Participant-oriented: Stance	*it is important to*	8	24	8
Participant-oriented: Stance	*so I think it*	5	17	2
Participant-oriented: Stance	*we have to*	8	24	3
Participant-oriented: Stance	*why I think*	0	16	2

L1 Korean

For L1 Korean writers, only three unique production tendencies were uncovered, all from the research-oriented functional category. Given the general underuse of LBs among L1 Korean writers, it is unsurprising that two of the three unique production tendencies in Table 6 were related to relative underuse. For the overused item (*is very expensive*), concordance lines revealed that this LB was exclusively used to bring focus to the high cost of tuition students face when attending post-secondary institutions.

Table 6
Individual Item Overuse/Underuse by L1 Korean Writers

Category	Bundle	Chinese	Japanese	Korean
Research-oriented: Purpose	*for us to*	34	12	0
Research-oriented: Quantification	*is very expensive*	1	1	14
Research-oriented: Quantification	*the most important*	22	14	2

Discussion

In contrast to most previous research, the current study avoided L1/L2 comparisons in favor of a contrastive interlanguage approach involving multiple L1 groups to better identify potential L1-related tendencies. This decision was largely driven by the growing recognition that L1/L2 comparisons suffer from a comparative fallacy that implies L2 users are somehow deficient in their language use (Hunston, 2002; Larsen-Freeman, 2014) and that using L1 discourse as a baseline hinders accurate descriptions of the L2 variety being analyzed (Bley-Vroman, 1983). Results of these inter-language (i.e., L2) comparisons revealed numerous differences in LB use by each L1 that suggest varying approaches to academic English writing.

L1 Chinese

L1 Chinese writers fell between Japanese and Korean in terms of total LB tokens. This contrasts with Appel and Murray (2020), which identified substantially more LBs in the L1 Chinese ESL corpus than in either the L1 Arabic or French corpora used for comparison. One potential reason for this discrepancy is the choice of L1 groups in each study. For example, Appel and Murray suggested that the relatively high number of LBs in L1 Chinese ESL writing may have resulted from emphasis given to collectivist thinking in Chinese culture. Given the more similar cultural basis of the three groups of L1 writers in the current study, this distinction may have become less apparent (see Hofstede et al., 2010).

Although the inclusion of more culturally comparable writer groups may have reduced the uniqueness of L1 Chinese production in this regard, individual LB patterns did indicate a uniquely collectivist approach among these writers. For example, aside from two overused bundles containing the first-person singular (*as far as I, in my opinion*), overuse of first-person plural (*for us to*, *we all know*) and underuse of first-person singular (*I want to, I think that*) were found. These findings, particularly overuse of first-person plural, may suggest a preference for LBs expressing 'in group' membership perhaps indicating how these writers view their position within a collectivist society. This argumentative approach positions writer and reader as part of the same collective, with those who disagree inherently viewed as outsiders, evidenced in the examples listed below where the writers seem to pursue inherent agreement with their position:

> *<W_CHN_PTJO_003>* ***We all know*** *that food made by ourselves tastes more delicious.*

> *<W_CHN_PTJO_005> As **we all know**, students have a lot of free time to manage.*
>
> *<W_CHN_PTJO_014> As **we all know**, college tuition is not a small sum.*

Further support for this position is found in Liardet (2018), where *we all know* was also identified as commonly used by Chinese EFL writers. Liardet situates this pattern as a subjective contracting evaluation that separates 'in group' agreement from 'out group' dissent. That L1 Chinese EFL writers are the lone group in this study from a communist society, wherein greater encouragement may be given to group membership, is likely an influential factor in the writing produced and LBs identified. However, as additional factors may also be at play, not all tendencies should be seen as a result of cultural influences, and factors such as pedagogic materials and instructional approach need to be analyzed in future research on this topic (see below).

The high number of text-transition bundles in the L1 Chinese corpus is in line with Leedham and Cai (2013), who examined production patterns for linking adverbials in L1 Chinese EFL writing. Of the six 3-4-word items they found to be overused when compared to L1 English writers, three were also overused relative to other L2 English users in the present study (*at the same time*, *last but not [least]*, and *what's more*), suggesting they may be regularly repeated patterns among L1 Chinese EFL writers.

Leedham and Cai (2013) ascribe this to teaching materials and an emphasis on rote learning within the Chinese education system. They also note that pedagogic approaches in mainland China typically fail to distinguish register, resulting in frequent use of less academic phrases, such as *what's more*, a bundle repeatedly identified as a distinctive feature of L1 Chinese L2 English writing (e.g., Appel & Szeib, 2018; Lee & Chen, 2009; Leedham & Cai, 2013). Given that the current study also found overuse for this item, pedagogic interventions may be necessary to reduce usage, with greater attention given to the importance of genre and register differences.

L1 Japanese

L1 Japanese writers had the highest overall number of LBs, with the largely similar nature of their writing suggesting collectivism, yet also frequent use of singular first-person pronouns, signifying individualism. Nam (2016) also identified relative overuse of first-person pronouns among L1 Japanese participants when comparing L1 Korean and Japanese EFL writers. This dichotomy between group and self may be accounted for somewhat by

Hofstede et al. (2010), who position Japan at a midway between collectivism and individualism. Again, however, more research, including analyses of pedagogic materials, will be needed to better understand these findings. For example, Northbrook and Conklin (2018) uncovered a high frequency of LBs featuring pronouns in textbooks used in Japanese junior high schools, arguing that frequent exposure could lead to subsequent language use. Thus, the high number of overused bundles incorporating first-person pronouns identified in the current study, as well as Nam, could be a lasting influence of pedagogic materials.

L1 Japanese were also the most frequent users of longer bundles, many of which incorporated portions of the essay prompt, with the full prompt occurring 72 times (compared with 24 and 23 occurrences in the L1 Chinese and L1 Korean corpora, respectively). This contrasts with the findings of Appel and Murray (2020), whose analysis of three L1 groups (Arabic, Chinese, and French) revealed L1 Chinese as the most frequent users of longer LBs, including those drawn from the essay prompts. Again, this difference may be a result of including more socially similar L1 groups in the present study. Granted partial use of the essay prompt is not in itself bad practice–Wray and Pegg (2009) note that it is also common practice among L1 writers–verbatim copying of the entire prompt appears to be a feature unique to the L1 Japanese context (at least when compared to L1 Chinese and Korean writers of English in this study).

Two longer prompt related phrases from the L1 Japanese group's top 10 bundles further highlight this feature: *I agree with the statement* was identified as overused by L1 Japanese writers (33 occurrences), with only one instance in the L1 Chinese corpus and entirely absent from the L1 Korean corpus, and *I agree with this* (25 times in the L1 Japanese corpus) occurred only one time in the L1 Chinese and Korean corpora. Although it is difficult to pinpoint the root cause of these tendencies, together with the highly frequent use of the various portions of the essay prompt, previous language teaching pedagogy may again be at play. The findings suggest that L1 Japanese students are being told to clearly signal their opinion with reference to the essay question, and are seemingly being provided phrases for doing so. However, further research involving L1 Japanese learners and the language teaching they commonly receive would be needed to confirm this.

L1 Korean

L1 Korean EFL writers were the least frequent users of LBs overall, apparent primarily through lower use of text- and participant-oriented bundles.

Comparatively low frequencies were also apparent in individual LBs, indicating less overlap, and greater language diversity.

With formulaic language a prevalent feature of academic writing (Hyland, 2008), lower frequency may signal reduced adherence to genre/register norms. However, because all writers were assessed to be at a similar proficiency level, further studies are needed to more closely examine this issue. Other potential explanations include a greater willingness among L1 Korean EFL writers to express themselves in non-standard ways and potentially higher lexical diversity.

L1 Korean EFL writers were especially infrequent users of participant-stance bundles. Jaworska et al. (2015) examined the use of stance expressions (labelled in the present study as 'participant-stance oriented') by L1 English and L1 German groups in argumentative essays in German, ascribing the L1 English group's greater use of impersonal and cautious language when expressing stance not to L1/L2 status, but to the transference of L1 rhetorical conventions. From this perspective, certain cultures place greater responsibility on writers to make text organization explicit, whereas others (including Chinese, Japanese, and Korean) place onus for understanding with the reader (see Leńko-Szymańska, 2008).

If, as claimed, all three groups in this study follow the same convention regarding text organization, the differences identified here in terms of the L1 Korean EFL writers' relatively low use of both transition and participant stance bundles may again be related to pedagogy. Leńko-Szymańska (2008) examined linking expressions among various L1 groups, finding significantly different levels of use between L1 groups classified as belonging to the same writing tradition, yet none between groups from different writing traditions, ascribing these differences in part to home country pedagogy. Here again, however, these conclusions are tentative and further research, which includes corpora of pedagogic materials used in each home country, is needed to more closely analyze each potential source of production differences.

Implications

Although not all identified production differences suggest a need for pedagogic intervention, L1-specific targeted instruction may be beneficial in improving the appropriateness of each group's academic English. Furthermore, the commonalities that were discovered suggest a combination of targeted (i.e., L1-specific) and general (i.e., universal to all L2 English users) instruction may prove beneficial.

In the case of L1 Chinese EFL writers, overuse of LBs less appropriate to the academic written register (e.g., *last but not [least]*, *what's more*) may require pedagogic interventions. Furthermore, as noted above, this group's use of stance bundles tended towards signaling a more subjective contracting evaluation (i.e., one suggesting in-group agreement and rejecting out-group dissent), which may be less appropriate in academic writing. Thus, focused instruction in register-appropriate transition phrases and stance expressions should be incorporated into the pedagogic approach for these students. Again, however, our findings were based on a relatively limited data set and future research assessing how well these findings apply to the target population at large is needed.

L1 Japanese EFL learners tended to overuse a narrow range of stance bundles and high number of bundles with first-person pronouns. Phrases including *I think* feature prominently and have been found with high frequency in previous studies of L1 Japanese corpora (e.g., Kobayashi, 2009; McCrostie, 2008), yet also among other L1 groups (e.g., Petch-Tyson, 1998; Ringbom, 1998). Though the present study found overuse in comparison to the other two L1 groups, more research may be needed to establish if it is indeed an L1-specific phenomenon or a more widespread issue among L2 learners. Kobayashi (2009) does however attribute the high frequency among L1 Japanese learners to L1 transfer. Contrastingly, Fordyce (2014) states that the 'more difficult' stance phrases typically make use of modal verbs, a structure he argues is problematic for L1 Japanese learners as it does not exist in the L1. Thus, there may be a need for focused instruction and guided practice making use of these sequences. L1 Japanese EFL writers' heavy reliance on personal pronouns may require similar treatment; if this is indeed a feature L1 Japanese learners have acquired through home country pedagogic materials, explicit instruction to reverse this tendency will be needed and an effort to revise these materials may be necessary.

L1 Korean EFL writing indicated general underuse of LB types and tokens, suggesting greater intra-group variance and potential deviance from standard academic written norms. This underuse was especially common with bundles used to signal transition and stance, both of which one might expect to be clearly marked within an argumentative essay. Combined, these factors indicate a potential lack of awareness regarding meeting register expectations through appropriate LB use. To highlight the importance of transitions, students could complete activities comparing texts lacking sufficient transition signals with those making effective use. To avoid the issue of register inappropriate use of transitions (e.g., *what's more*) identi-

fied in the L1 Chinese group, attention should be paid to including academic expressions. Fill-in-the blank activities, where students select register and context appropriate transitions, could also be used.

A second area with potential pedagogic implications relates to the influence of teaching materials within the L1 context. For L1 Chinese, this was seen in the overuse of text-transition bundles, similar to those identified by Leedham and Cai (2013), who suggest that sample texts and vocabulary lists typically provided to students in China may be partly to blame. Similarly, overuse of LBs featuring pronouns identified among L1 Japanese writers may result from teaching materials (Northbrook & Conklin, 2018). Thus, the influence of pedagogic materials may be a factor that merits closer examination in future studies. If such pedagogic factors are found to cause L1 specific over/underuse, remedying the materials at fault would be more beneficial than attempting to subsequently counter the symptoms created.

Finally, in terms of methodological implications, an aspect addressed by Appel and Murray (2020) and further developed here (i.e., the value of not limiting extracted sequence length to the common 4-word length, but instead including all bundles which fulfill the identification criteria applied), deserves mention due to the impact this can have on results. Appel and Murray went beyond the typically applied 4-word bundle limit to include all items from three to seven words in length, though the current study opted to remove the upper word limit altogether. Although restricting sequence length may help maintain a focus on the most frequent patterns, if, as here, the purpose is to examine and compare actual language use across corpora, employing identification criteria which allow the full range of these differences to surface seems critical. Doing so allows for a more complete picture by looking beyond the 4-word structures commonly sought which could often be more accurately viewed as single, extended items. For example, in the present study, this approach was used to help identify frequent and verbatim use of the essay prompt by L1 Japanese EFL writers, a factor which may otherwise have gone unnoticed.

Limitations and Future Research

Limitations in the current study arise from two main areas which future researchers should attempt to address. First, although efforts were made to control for proficiency across L1 groups, the essays included in the ICNALE use automated measures to classify writers into each proficiency band. Thus, more strictly controlled proficiency measures may be needed to more adequately control for the influence of proficiency differences on findings.

The second main limitation concerns the lack of understanding regarding root causes of identified production differences. As an exploratory study, we aimed to identify L1 differences to make tentative suggestions regarding why such tendencies were exhibited. Given the implications of such features as overuse of register-inappropriate language in terms of how a writer is assessed, further studies that look more closely at potential root causes for these tendencies are clearly needed in order to redress them. In-depth examinations of pedagogic materials used in each country, collections of L1 writing from each group, and potential translation equivalents, could all prove beneficial.

Conclusion

Findings from this study contribute to the growing body of research suggesting that particular L1 groups produce L2 language in significantly distinct ways. Production differences were found to occur at various levels of analysis, including overall LB type/token counts, functional category, and individual item, indicating areas where each L1 group may benefit from focused pedagogic interventions. Various factors were proposed as influencing these distinct production tendencies, including cultural elements, L1 transfer, and language teaching pedagogy. However, the influence of each factor on production remains unclear, meaning further research is necessary to bring greater clarity to the issues at hand.

Notes

1. The ICNALE uses scores from high-stakes proficiency (e.g., TOEFL, TOEIC) and vocabulary size tests to map L2 writers' proficiency ratings onto CEFR levels.

Randy Appel is an associate professor in the Global Education Center at Waseda University. His research interests include corpus approaches to linguistic inquiry, L2 English academic writing, and the use of formulaic language by L1 and L2 users.

Lewis Murray is an associate professor in the Foreign Language Education department at Kanazawa University. His research interests relate primarily to L2 learners' acquisition of formulaic language in relation to academic writing.

References

Adel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes, 31*(2), 81–92. https://doi.org/10.1016/j.esp.2011.08.004

Allen, D. (2011). Lexical bundles in learner writing: An analysis of formulaic language in the ALESS learner corpus. *Komaba Journal of English Education, 1*, 105–127. https://park.itc.u-tokyo.ac.jp/eigo/KJEE/001/105-127.pdf

Appel, R., & Murray, L. (2020). L1 differences in L2 English academic writing: A lexical bundles analysis. *Journal of English for Academic Purposes, 46*, 1–15. https://doi.org/10.1016/j.jeap.2020.100873

Appel, R., & Szeib, A. (2018). Linking adverbials in L2 English academic writing: L1 related differences. *System, 78*, 115–129. https://doi.org/10.1016/j.system.2018.08.008

Appel, R., & Trofimovich, P. (2017). Transitional probability predicts native and non-native use of formulaic sequences. *International Journal of Applied Linguistics, 27*(1), 24–43. https://doi.org/10.1111/ijal.12100

Appel, R., & Wood, D. (2016). Recurrent word combinations in EAP test-taker writing: Differences between high- and low-proficiency levels. *Language Assessment Quarterly, 13*(1), 55–71. https://doi.org/10.1080/15434303.2015.1126718

Biber, B., Conrad, S., & Cortes, V. (2004). If you look at …: Lexical bundles in university teaching and textbooks. *Applied Linguistics, 25*(3), 371–405. https://doi.org/10.1093/applin/25.3.37

Biber, B., Conrad, S., & Reppen, R. (1999). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press. https://doi.org/10.1017/CBO9780511804489

Bley-Vroman, R. (1983). The comparative fallacy in interlanguage studies: The case of systematicity, *Language Learning*, *33*(1), 1–17. https://doi.org/10.1111/j.1467-1770.1983.tb00983

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. https://doi.org/10.1016/j.jeap.2017.10.008

Chen, A. C.-H. (2019). Assessing phraseological development in word sequences of variable lengths in second language texts using directional association measures. *Language Learning, 69*(2), 1–38. https://doi.org/10.1111/lang.12340

Chen, Y., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning and Technology, 14*(2), 30–49. https://doi.org/10125/44213

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes, 23*(4), 397–423. https://doi.org/10.1016/j.esp.2003.12.001

Durrant, P. (2017). Lexical bundles and disciplinary variation in university students' writing: Mapping the territories. *Applied Linguistics, 38*(2), 165–193. https://doi.org/10.1093/applin/amv011

Fordyce, K. (2014). The differential effects of explicit and implicit instruction on EFL learners' use of epistemic stance. *Applied Linguistics, 35*(1), 6–28. https://doi.org/10.1093/applin/ams076

Granger, S., & Thewissen, J. (2005). *Towards a reconciliation of "can do" and "can't do" approach to language assessment.* [Paper presentation]. The Second Annual Conference of the European Association of Language Testing and Assessment, Voss, Norway.

Hofstede, G., Hofstede, G., & Minkov, M. (2010). *Cultures and organizations: Software of the mind* (Rev. 3rd ed.). McGraw-Hill.

Hunston, S. (2002). Pattern grammar, language teaching, and linguistic variation: Applications of a corpus-driven grammar. In R. Reppen, S. M. Fitzmaurice, & D. Biber (Eds.), *Using corpora to explore linguistic variation* (pp. 167–183). John Benjamins. https://doi.org/10.1075/scl.9.11hun

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes, 27*, 4–21. https://doi.org/10.1016/j.esp.2007.06.001

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. Routledge. https://doi.org/10.4324/9780203935927

Jaworska, S., Krummes, C., & Ensslin, A. (2015). Formulaic sequences in native and non-native argumentative writing in German. *International Journal of Corpus Linguistics, 20*(4), 500–525. https://doi.org/10.1075/ijcl.20.4.04jaw

Karabacak, E., & Qin, J. (2013). Comparison of lexical bundles used by Turkish, Chinese, and American university students. *Procedia - Social and Behavioral Sciences, 70,* 622–628. https://doi.org/10.1016/j.sbspro.2013.01.101

Kobayashi, Y. (2009). Extracting characteristics of English essays written by Japanese EFL learners based on the NS/NNS text categorization model. *Proceedings of Jinmonkon Symposium Series of Information Processing Society of Japan, 16,* 261–268.

Kuiper, K. (1996). *Smooth talkers: The linguistic performance of auctioneers and sportscasters*. Lawrence Erlbaum.

Larsen-Freeman, D. (2014). Another step to be taken: Rethinking the endpoint of the interlanguage continuum. In Z. H. Han & E. Tarone (Eds.), *Interlanguage: Forty years later* (pp. 203–220). John Benjamins. https://doi.org/10.1075/lllt.39.11ch9

Lee, D., & Chen, S. (2009). Making a bigger deal of the smaller words: Function words and other key items in research writing by Chinese learners. *Journal of Second Language Writing, 18*(4), 149–165. https://doi.org/10.1016/j.jslw.2009.07.003

Leedham, M., & Cai, G. (2013). Besides…on the other hand: Using a corpus approach to explore the influence of teaching materials on Chinese students' use of linking adverbials. *Journal of Second Language Writing, 22*(4), 374–389. https://doi.org/10.1016/j.jslw.2013.07.002

Leńko-Szymańska, A. (2008). Non-native or non-expert? The use of connectors in native and foreign language learners' texts. *Aile: Acquisition et Interaction en Langue Étrangère, 27*, 91–108. https://doi.org/10.4000/aile.4213

Liardet, C. (2018). 'As we all know': Examining Chinese EFL learners' use of interpersonal grammatical metaphor in academic writing. *English for Specific Purposes, 50*, 64–80. https://doi.org/10.1016/j.esp.2017.11.005

Lu, X., & Deng, J. (2019). With the rapid development: A contrastive analysis of lexical bundles in dissertation abstracts by Chinese and L1 English doctoral students. *Journal of English for Academic Purposes, 39*, 21–36. https://doi.org/10.1016/j.jeap.2019.03.008

McCrostie, J. (2008). Writer visibility in EFL learner academic writing: A corpus-based study. *ICAME Journal, 32,* 97–114. http://icame.uib.no/ij32/ij32_97_114.pdf

Nam, D. (2016). Lexical bundle structures in argumentative writing of Korean and Japanese learners of English. *Studies in Foreign Language Education, 30*(3), 1–24. https://doi.org/10.16933/sfle.2016.30.3.1

Northbrook, J., & Conklin, K. (2018). "What are you talking about?" An analysis of lexical bundles in Japanese junior high school textbooks. *International Journal of Corpus Linguistics, 23*(3), 311–334. https://doi.org/10.1075/ijcl.16024.nor

Ortega, L. (2011, June). *New trends in SLA research: Theories, methods, ethics*. Hsinchu, Taiwan: Invited lecture at the National Tsing Hua University.

Paquot, M. (2017). L1 frequency in foreign language acquisition. Recurrent word combinations in French and Spanish EFL learner writing. *Second Language Research, 33*(1), 12–32. https://doi.org/10.1177/0267658315620265

Petch-Tyson, S. (1998). Writer/reader visibility in EFL written discourse. In S. Granger (Ed.), *Learner English on computer* (pp. 107–118). Routledge. https://doi.org/10.4324/9781315841342-8

Ringbom, H. (1998). Vocabulary frequencies in advanced learner English: A cross-linguistic approach. In S. Granger (Ed.), *Learner English on computer* (pp. 41–52). Routledge. https://doi.org/10.4324/9781315841342-3

Schmitt, N. (Ed.). (2004). *Formulaic sequences: Acquisition, processing and use.* John Benjamins. https://doi.org/10.1075/lllt.9

Shin, Y. (2019). Do native writers always have a head start over nonnative writers? The use of lexical bundles in college students' essays. *Journal of English for Academic Purposes, 40*, 1–14. https://doi.org/10.1016/j.jeap.2019.04.004

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press. https://doi.org/10.1017/CBO9780511519772

Wray, A., & Pegg, C. (2009). The effect of memorized learning on the writing scores of Chinese IELTS test takers. *IELTS Report, 9*, 191–216. https://www.ielts.org/-/media/research-reports/ielts_rr_volume09_report5.ashx