



Interactions in EFL argumentative writing: effects of topic, L1 background, and L2 proficiency on interactional metadiscourse

Hyung-Jo Yoon¹

Published online: 3 September 2020
© Springer Nature B.V. 2020

Abstract

This study examined how students of English as a foreign language (EFL) with different first language (L1) backgrounds use interactional metadiscourse markers in argumentative writing. Specifically, to explore unique patterns of metadiscourse features that reflect context and development, the essays written by Chinese, Japanese, and Korean EFL students at three proficiency levels were analyzed for topic, L1 background, and L2 proficiency. For a comprehensive analysis of 1986 essays, I used a natural language processing tool that generates quantity scores for Hyland's (2005) metadiscourse categories (i.e., hedges, boosters, attitude markers, self-mentions, reader pronouns, and directives). The results showed notable differences in the students' use of metadiscourse features across topics, and significant variation was also found across different L1 groups. However, their use of interactional metadiscourse did not differ by L2 proficiency. A post hoc analysis of a parallel native-speaker corpus further revealed EFL students' underuse of hedges and overuse of reader pronouns. Findings are discussed in terms of academic writing instruction, writing prompt development, and L2 learner categorization.

Keywords Authorial voice · Corpus-based research · Difference across L1 groups · EFL writing · Interactional metadiscourse · Topic effect · Second language writing

Introduction

Over the years, researchers have shown an increased interest in exploring student writers' use of metadiscourse markers, with a focus on first language (L1) and proficiency effects (e.g., Aull, 2019; Bax, Nakatsuhara, & Waller, 2019; Ho & Li, 2018; Hong & Cao, 2014; Lee & Deakin, 2016). The concept of metadiscourse has been

✉ Hyung-Jo Yoon
hyungjo.yoon@csun.edu

¹ Department of Linguistics/TESL, California State University, Northridge, 18111 Nordhoff St., Northridge, CA 91330, USA

defined and understood differently across scholars; in this study, metadiscourse is viewed as “linguistic resources used to organize a discourse or the writer’s stance towards either its content or the reader” (Hyland & Tse, 2004, p. 157). It particularly focuses on the interactional dimension of metadiscourse that allows writers to convey their ideas more effectively to their audience (Hyland, 2005). Previous second language (L2) studies on interactional metadiscourse have demonstrated that learner essays, when compared with those by native English-speaking (NS) counterparts, tend to be characterized by the misuse and underuse of metadiscourse features (e.g., Çandarlı, Bayyurt, & Martı, 2015; Hyland & Milton, 1997; Lee & Deakin, 2016). Other studies have also examined the use of interactional metadiscourse markers in relation to learners’ L1 background (e.g., Hong & Cao, 2014) and L2 proficiency (e.g., Bax et al., 2019), and their findings have commonly pointed to the importance of increased pedagogical attention to interactional metadiscourse for L2 writers (detailed review of previous findings offered in the literature review section).

This study is grounded in Hyland’s (2005) interactional metadiscourse model which consists of two dimensions, namely, *stance* and *engagement*. Stance involves writer-oriented resources of interaction that writers would use when presenting their opinions and attitudes toward a proposition. In Hyland’s model, stance includes the categories of *hedges*, *boosters*, *attitude markers*, and *self-mentions*. Hedges are linguistic resources that express writers’ uncertainty about a proposition and accept the possibility of alternative ideas (e.g., *arguably*, *maybe*, *usually*), while boosters are to convey writers’ commitment and confidence in their argument (e.g., *every single*, *evidently*, *obviously*). Attitude markers show writers’ emotions and affective attitude towards a proposition (e.g., *amazingly*, *cruelly*, *unusually*). Self-mentions refer to explicit authorial presence in the text, which is realized mostly through first person pronouns (Hyland, 2001).

The other side of the framework, *engagement*, involves the writer’s recognition of readers’ presence and willingness to include them as legitimate discourse participants. The categories of engagement include *reader pronouns*, *directives*, *questions*, *knowledge reference*, and *personal asides*. Reader pronouns involve the use of second person pronouns and, in some contexts, first person plural pronouns (inclusive *we*). Directives are intended to direct readers to perform some acts mainly through the use of imperative constructions and obligation modals (Hyland, 2002a). From the engagement categories, this study examines only reader pronouns and directives because the remaining categories (questions, knowledge reference, and personal asides) have been found to occur minimally in timed argumentative writing (Hyland, 2005; Zhao, 2017).

While Hyland’s (2005) comprehensive model has gained much attention for empirical research, previous studies tended to focus on only a few categories (mostly hedges and boosters) and quantified occurrences of metadiscourse markers based on their subjective judgments that are often different from other studies. Also, potentially due to the labor-intensive nature of manually identifying target markers, the number of learner essays used for the exploration of metadiscourse has been relatively small. In an effort to see the developmental trajectory of how L2 writers express their attitudes toward a proposition and awareness of an audience, the current study involves a comprehensive quantitative analysis of 1986 learner essays

using an automated processing tool that computes quantity scores of each metadiscourse category (Yoon, 2017a). The focus of this analysis is on identifying notable patterns of interactional metadiscourse by two learner variables (L1 background and L2 proficiency) and one task variable (writing topic). This study specifically attends to East Asian EFL students with different L1 backgrounds (Chinese, Japanese, and Korean) to reveal potential L1 background effects. These learner groups are targeted because of the large proportion of English language learners speaking these languages; students from China constitute 33.7% of the international student population studying in the United States, and those from China, Japan, and South Korea add up to 40.2% (Institute of International Education, 2019). Results from the learner data are then compared with those from native English speaker essays, a suitable target for EFL writers. Writing topic, as one of the task characteristics, has been explored in previous L2 writing research (e.g., Yoon, 2017b; Yang & Kim, 2020; Yang, Lu, & Weigle, 2015), but scholars have mostly attended to topic effects on linguistic features of syntactic complexity, accuracy, and fluency. Addressing these gaps in the literature, this corpus-based study aims to achieve a better understanding of L2 writers' use of interactional metadiscourse markers in written argumentation and offer implications for writing instruction and assessment.

Interactional metadiscourse in L2 writing research

Argumentative writing has been the most typical genre for adult L2 writers due to its prevalent use in educational and testing contexts (Qin & Karabacak, 2010; Wolfe, 2011). Thus, despite the complex nature of establishing one's position for convincing arguments, adult L2 writers are likely to have built task schemas and genre awareness for written argumentation. In the context of argumentative and research writing, previous studies have endeavored to identify distinct patterns of metadiscourse use between L1 and L2 writers (e.g., Çandarlı et al., 2015; Hyland & Milton, 1997; Lee & Deakin, 2016), and between L2 writers with varied L1 backgrounds (e.g., Hong & Cao, 2014; Lee & Casal, 2014). Using a contrastive perspective, these studies have commonly informed us that L2 learners have difficulties attaining native-like command of interactional metadiscourse in writing. For example, an early study by Hyland and Milton (1997) compared Chinese L2 learners' essays with those by L1 writers with regard to hedge and booster expressions. They found L2 learners' overuse of boosters and underuse of hedges, which led to Chinese learners' stronger commitments to their assertions in writing. Similarly, Lee and Deakin (2016) examined the three corpora of 25 high-rated English as a second language (ESL) essays, 25 low-rated ESL essays (by undergraduate students speaking Chinese as their L1), and 25 high-rated L1 essays (by NS undergraduate students) in terms of their inclusion of interactional metadiscourse markers. The results of Lee and Deakin showed that the high-rated ESL and L1 essays contained higher numbers of hedges than the low-rated ESL essays. Their findings also indicated that the ESL essays commonly contained much fewer self-mentions than the L1 essays, while occurrences of boosters and attitude makers did not differ significantly between the corpora.

L2 scholars have interpreted distinct patterns of metadiscourse features between L1 and L2 writing, with a focus on cultural and contextual factors, as well as awareness of academic conventions (e.g., Hyland & Milton, 1997; Lee & Deakin, 2016; Li & Warton, 2012). For the overall trend of fewer types of metadiscourse markers in L2 writing, it has been argued that L2 learners' limited repertoire of metadiscourse items lead them to overuse a narrow range of metadiscourse items repeatedly. It may also be their limited pragmatic competence that causes L2 writers' difficulties in using metadiscourse appropriately across registers (Hong & Cao, 2014; Hyland & Milton, 1997). That is, some L2 learners, even with their knowledge of a large number of metadiscourse markers, may still find it challenging to use interactional metadiscourse properly in their writing due to potential differences in their expected, conventional use of interactional resources between L1 writing and target writing contexts (Lee & Casal, 2014). Additionally, the finding that L2 learners tend not to use adequate numbers of hedging expressions in their writing has been regarded as an outcome of their inaccurate concern that the inclusion of hedges would weaken the convincing force of their arguments. Their limited use of self-mentions and reader pronouns has also been attributed to their learning experience that informal language should be avoided in English academic writing. From a cultural perspective, Chinese student writers have been found to use fewer self-mentions in constructing their English texts than L1 counterparts, and it has been interpreted as East Asian students' reluctance to express authorial presence in academic writing due to their collectivistic backgrounds (Bloch & Chi, 1995; Hyland, 2002b).

As illustrated above, the majority of previous research into interactional metadiscourse has focused on how L2 writing differs from L1 writing that serves as benchmark data, and notable differences between the two groups have been recognized as areas for pedagogical attention. While this line of research has advanced our understanding of interactional metadiscourse features in L2 writing, it should be noted that L2 learners are not a uniform group to be contrasted with L1 writers and their use of metadiscourse features may differ by L1 background. One attempt to address this issue was a corpus-based study by Hong and Cao (2014). They examined occurrences of interactional metadiscourse features in English as a foreign language (EFL) learner writing across three different L1s (i.e., Chinese, Polish, and Spanish). Their results showed no significant difference in the number of hedges among the learner groups but the essays by Chinese learners contained more boosters than the other two groups. This finding apparently coincides with earlier findings that Chinese writers would make assertive claims with reduced hedges and more boosters (Hyland & Milton, 1997), but it has often been considered evidence of how East Asian learners of English would use interactional resources, mainly due to the lack of research on interactional metadiscourse involving East Asian learners with different L1 backgrounds. Nevertheless, a few L2 studies endeavored to explore the writing of East Asian learners with different L1 backgrounds and in fact found some distinct patterns of language use among different L1 groups (e.g., Hinds, 1990; Hinkel, 2003, 2009). Based on the ethnic and linguistic composition of the East Asian English learner population (Williams, 2017), these studies tended to target Chinese, Japanese, and Korean speakers.

Interactional metadiscourse at different proficiency levels

Much attention has been given to the question of how L2 writers at different proficiency levels express their stance towards a proposition or the reader, and making arguments with a clear awareness of the reader's presence has become one of the areas of L2 writing development, as specified by CEFR descriptors for writing (Council of Europe, 2018). If the appropriate and varied use of interactional markers in written discourse is indicative of L2 writing development, we may need to better understand specific patterns of interactional marker use by L2 learners at different proficiency levels and to give insight into how L2 writing instruction can facilitate their development. A few studies have explored this issue of how learners' use of interactional metadiscourse markers varies by proficiency (e.g., Bax et al., 2019; Cheung & Low, 2019; Intaraprawat & Steffensen, 1995). For example, a recent study by Bax et al. (2019) examined a corpus of 900 essays composed by L2 learners at three different proficiency levels (CEFR B2, C1, C2 levels). Their findings showed that the essays by higher-level learners tended to display a greater variety of interactional markers than those by lower-level learners, while total numbers of markers included in the higher-level essays were fewer than those in the lower-level essays. Interestingly, this pattern of more types and fewer tokens with increased L2 proficiency was found for each of the stance categories targeted in their study (i.e., hedges, boosters, and attitude markers; except for self-mentions that only involve a few types of first person pronouns).

Another method employed in studies on interactional metadiscourse patterns across proficiency levels is to compare metadiscourse features in high quality essays with those in low quality essays (e.g., Cheung & Low, 2019; Ho & Li, 2018; Intaraprawat & Steffensen, 1995; Lee & Deakin, 2016). With the assumption that external measures of proficiency do not always correlate with writing quality, studies using this method have aimed to offer insight into the contributive role of metadiscourse markers for essay quality. Intaraprawat and Steffensen (1995), for example, examined the use of interactional metadiscourse markers (hedges, boosters, and attitude markers) in high-rated and low-rated essays, and they suggested greater occurrences of all metadiscourse categories in the high-rated essays than in low-rated ones. However, using a similar method of comparing high-graded and low-graded argumentative essays, Lee and Deakin (2016) found that the high-graded essays were similar to the low-graded essays in terms of boosters, attitude markers, and engagement markers. One notable difference between the two studies is that the participants of Lee and Deakin (2016) are all Chinese ESL students, while those of Intaraprawat and Steffensen (1995) are ESL students with a wide range of L1 backgrounds (e.g., Arabic, French, Japanese, Korean, and Spanish); it is plausible that their mixed findings might have arisen from the interaction between the two learner variables (i.e., L2 proficiency and L1 background). Thus, to disentangle the effect of L2 proficiency on interactional patterns in writing from that of L1 background, it may be necessary to explore a corpus of essays written by L2 learners from different L1 backgrounds who have been identified to be at different proficiency levels.

Topic effects on interactional metadiscourse

In the field of L2 writing research, there has been much research on the effect of topic on textual features (e.g., Yoon, 2017b; Yang & Kim, 2020; Yang et al., 2015). Specifically, as one of the main writing task variables, topic has been found to influence features of syntactic complexity (e.g., Hinkel, 2009; Yang et al., 2015) and lexical sophistication (e.g., Reid, 1990; Yang & Kim, 2020). Findings of topic effect research have informed us of how to interpret the contribution of syntactic and/or lexical features to writing quality validly in consideration of topic, which is a potential mediating factor. These studies have been motivated by clear evidence that some syntactic and lexical features are good indicators of writing proficiency and development (e.g., Bulté & Housen, 2014; Crossley & McNamara, 2014), and recent studies have begun to explore interactional metadiscourse as one of the areas that would be indicative of proficiency (e.g., Aull & Lancaster, 2014; Bax et al., 2019) and contribute to writing quality (e.g., Yoon, 2017a; Zhao, 2017), but there has been no empirical research that examines topic effects on interactional metadiscourse in L2 writing comprehensively.

Huot's (1990) review of writing prompts noted that the three feature categories—(a) discourse mode, (b) rhetorical specification (writing purpose and audience specified), and (c) the wording and structure of writing prompts—are likely to influence learners' writing performance, but previous findings of the relationships between these prompt features and writing performance have been inconclusive. Also, relatively little attention given to the wording of prompts indicates the need to explore potential wording effects, together with those of other macro-level prompt features, on L2 learners' language use. This study addresses the need for a large-scale quantitative analysis of interactional metadiscourse, with a focus on a task variable (writing topic) and two learner variables (L1 background and L2 proficiency). Given the lack of theoretical explanations of East Asian students' different use of interactional metadiscourse, this study is of an exploratory nature. The present research is guided by the following research questions:

1. How does writing topic affect EFL students' use of interactional metadiscourse in argumentative writing?
2. How do EFL students with different first language backgrounds use interactional metadiscourse in argumentative writing?
3. How do EFL students at different second language proficiency levels use interactional metadiscourse in argumentative writing?
4. How does EFL students' argumentative writing differ from that of NS writers in interactional metadiscourse use?

Method

Corpus data

For this study, I used the International Corpus Network of Asian Learners of English (ICNALE) (Ishikawa, 2013) that includes timed argumentative essays collected from college-level Asian EFL students from 10 countries (5400 essays) as well as native English speakers (400 essays). The EFL students have been classified into one of the adjusted CEFR levels (i.e., A2: waystage; B1.1: threshold, lower; B1.2: threshold, upper; and B2: vantage or higher), based on their performance on a proficiency test (e.g., TOEIC or TOEFL) or a vocabulary size test. Each student wrote two argumentative essays under time constraints; specifically, having access to the two prompts simultaneously, they had to complete two essays within a given time (up to 80 min for two essays). In the prompts, the students were asked to agree or disagree with the statements below:

1. It is important for college students to have a part-time job (hereafter, the Part-time topic).
2. Smoking should be completely banned at all the restaurants in the country (hereafter, the Smoking topic).

For its major quantitative analyses, this study used 1986 essays written by Chinese, Japanese, and Korean EFL students at the three proficiency levels (A2, B1.1, and B1.2). The essays at the B2 level were excluded from analysis because the number of students at this level was incomparably fewer than those at other levels (e.g., only 13 Chinese and 18 Japanese speaking students at B2). The learner essays were mostly written by college freshmen and sophomores (85.8%), and each of the L1 groups did not show a significant gender imbalance (203 female and 184 male Chinese speakers; 132 female and 250 male Japanese speakers; 140 female and 84 male Korean speakers). Identifying unique patterns of interactional metadiscourse features from the essays by these students will offer valuable implications for material development and writing instruction.

Additionally, I discuss to what extent East Asian EFL student writing is different from L1 writing with regard to interactional metadiscourse features by analyzing 400 L1 essays. It should be noted here that the L1-L2 comparison is not the focus of this study, thus limiting the use of L1 essay results to a post hoc analysis that follows detailed analyses of the EFL essays. Table 1 displays the overall size and the number of essays of each sub-corpus.

Instruments

This study employed the Authorial Voice Analyzer (AVA) that calculates normalized frequencies of Hyland's (2005) interactional metadiscourse categories (i.e., *hedge*, *booster*, *attitude marker*, *self-mention*, *reader pronoun*, and *directive*)

Table 1 Summary of the corpus data

L1	Proficiency	Number of essays	Number of words	Average words per essay
Chinese	A2	100	22,000	220.0
	B1.1	464	107,827	232.4
	B1.2	210	50,776	241.8
Japanese	A2	308	67,029	217.6
	B1.1	358	77,793	217.3
	B1.2	98	21,810	222.6
Korean	A2	150	32,541	216.9
	B1.1	122	26,272	215.3
	B1.2	176	39,178	222.6
English		400	89,064	222.7
Total		2386	534,290	223.9

Table 2 Interactional metadiscourse features from AVA

Feature	Token	Type	Normalized frequency
Hedge	✓	✓	Occurrences per 1000 words
Booster	✓	✓	
Attitude marker	✓	✓	
Self-mention	✓		
Reader pronoun	✓		
Directive	✓		

(Yoon, 2017a). Of these categories, AVA produces both token (total number of items) and type (number of unique items) values for hedges, boosters, and attitude markers (i.e., categories with a wide variety of items) to identify potentially different roles of the extensive and varied use of stance features. For each category, as presented in Table 2, AVA computes normalized values (i.e., occurrences per 1000 words), which have been widely employed in previous research into interactional metadiscourse (e.g., Aull & Lancaster, 2014; Hong & Cao, 2014; Hyland, 2005; Lee & Deakin, 2016).

AVA has been designed to cover a wide range of interactional metadiscourse markers and capture linguistic variability in word sequences through regular expressions. One example of booster devices is “without\s\w*\s?doubt*” in which “\s\w*\s?” indicates the possible insertion of an optional word. As a result, this item counts instances of *without doubt*, *without any doubt*, *without much doubt*, *without a doubt*, etc. An example of hedges is “(ilwe)\s\w*\s?\w*\s?assume” that includes “I” (i.e., or) and “\s\w*\s?\w*\s?” (i.e., optional insertion of one word or two words), and this item counts instances of *I assume*, *I would assume*, *I just simply assumed*, *we sort of assumed*, etc. Another example from the hedge list

is “(? < !inlun)(arguable|arguably)” that includes “? < !” (i.e., *not following*); this item counts occurrences of *arguable* and *arguably* as hedges but not *inarguable*, *inarguably*, or *unarguable*.

The hedge list of AVA includes a total of 164 expressions, and the booster list includes a total of 174 expressions. The attitude marker list built based on the existing lists of emotion and attitude words (Hu & Liu, 2004; Mohammad & Turney, 2013) contains 640 expressions. The categories of self-mentions and reader pronouns include personal pronouns: first person singular pronouns (*I, my, me, and mine*) for self-mentions, and first person plural (*we, our, us, ours*) and second person pronouns (*you, your, yours*) for reader pronouns. First person plural pronouns are categorized as reader pronouns, given their primary function as inclusive *we* in single-authored argumentative writing. For directives, AVA counts obligation modals (e.g., *should, ought, have to, need to, must*) and predicative necessity-related adjectives controlling a to-clause complement (e.g., *it is essential to, it is necessary to, it is obligatory to, it is vital to*) reported in Hyland (2002a). These predicative adjectives of necessity are all constructed with relevant regular expressions. For example, a directive item, *it is essential to*, is added to the list as “(it is|its)\s*\w*\s*?essential\s(for)?\s*\w*\s*\w*\s*\w*\s*to”, making it count varied expressions of this extraposed *to* clause (e.g., *it is essential to, it’s essential for him to, it’s sometimes essential for a college student to*, etc.). For directives, AVA also counts the number of imperative constructions that are identified by the Stanford parser (Klein & Manning, 2003).

This study examines a very large number of essays by EFL students at different proficiency levels. Normalized frequencies of *tokens* on the basis of 1000 words are used as target measures in this study. To validate these measures, Yoon and Römer (2020) hand coded 20 randomly selected MICUSP texts for hedge and booster features. After resolving all discrepancies through discussion, we checked the correlations of hand coded values with those from AVA. As a result, we found a correlation coefficient of .921 on hedges and .892 on boosters, offering evidence of AVA validly quantifying interactional metadiscourse markers. The full lists of the AVA items are available in the supplemental material of Yoon and Römer (2020).

Analysis

In exploring the main and interaction effects of three predictor variables (i.e., topic, L1 background, and L2 proficiency), I performed three-way ANOVAs with topic as a within-subjects variable and L1 background and L2 proficiency as between-subjects variables. In these analyses, normalized frequencies of metadiscourse markers were included as dependent variables. For self-mentions and reader pronouns, I used logarithmically transformed values to address their positively skewed distributions (Field, 2013). Each of the dependent variables showed descriptive results that satisfy assumptions for three-way ANOVAs (see Table 3 for descriptive statistics). In interpreting statistical significance, the alpha level adjusted by the Bonferroni correction ($.0083 = .05/6$) was adopted to decrease the possibility of Type I error. The magnitude of effects was estimated using partial eta squared (η_p^2), a valid effect size

Table 3 Mean and standard deviation values of interactional metadiscourse by topic, L1 background, and L2 proficiency

Topic	L1	Proficiency	Hedge		Booster		Attitude		Self (log)		Reader (log)		Directive	
			Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	
Part-time	Chinese	A2	13.01 (8.02)	26.76 (12.75)	15.84 (8.83)	1.31 (.19)	1.57 (.32)	12.17 (8.16)						
		B1.1	14.66 (9.06)	27.43 (11.12)	15.73 (8.49)	1.30 (.21)	1.65 (.33)	11.53 (7.02)						
		B1.2	16.65 (9.02)	27.28 (10.99)	15.23 (7.93)	1.28 (.19)	1.62 (.33)	10.38 (7.07)						
	Japanese	A2	20.63 (10.31)	28.62 (14.92)	18.59 (9.44)	1.56 (.27)	1.50 (.39)	17.16 (10.84)						
		B1.1	19.43 (9.37)	28.61 (13.02)	18.81 (8.52)	1.56 (.29)	1.49 (.39)	16.94 (9.67)						
		B1.2	18.87 (10.34)	27.38 (11.90)	21.30 (9.02)	1.59 (.29)	1.38 (.33)	15.63 (8.86)						
Smoking	Korean	A2	11.95 (9.63)	22.60 (13.45)	15.49 (8.98)	1.42 (.30)	1.46 (.31)	10.68 (9.60)						
		B1.1	13.99 (9.76)	20.92 (12.02)	13.37 (8.95)	1.52 (.31)	1.43 (.33)	12.22 (9.16)						
		B1.2	15.81 (9.41)	20.80 (10.75)	16.39 (8.79)	1.44 (.32)	1.37 (.36)	9.86 (6.99)						
	English	A2	25.55 (13.23)	25.23 (11.76)	18.13 (9.00)	1.46 (.32)	1.22 (.28)	10.23 (7.07)						
		B1.1	16.31 (9.09)	40.38 (15.32)	14.80 (8.09)	1.33 (.17)	1.47 (.24)	14.75 (8.95)						
		B1.2	17.67 (9.03)	39.73 (14.94)	16.50 (9.12)	1.31 (.18)	1.49 (.23)	13.67 (7.33)						
Japanese	A2	16.73 (8.73)	41.86 (15.57)	17.74 (7.96)	1.28 (.19)	1.44 (.23)	13.56 (7.14)							
		19.95 (11.50)	39.79 (13.85)	17.56 (9.69)	1.45 (.26)	1.33 (.25)	14.51 (8.34)							
		20.07 (10.92)	39.27 (15.56)	17.42 (8.79)	1.44 (.27)	1.32 (.26)	14.63 (8.04)							
	B1.1	18.65 (9.13)	44.57 (17.78)	17.91 (9.25)	1.41 (.24)	1.33 (.27)	16.48 (7.31)							
		9.61 (8.80)	34.76 (16.54)	19.45 (10.02)	1.34 (.25)	1.44 (.29)	15.29 (10.08)							
		13.41 (10.04)	34.66 (19.08)	18.80 (9.87)	1.42 (.32)	1.35 (.20)	14.76 (10.78)							
B1.2	14.36 (10.98)	34.88 (17.43)	17.80 (10.04)	1.34 (.22)	1.29 (.25)	14.32 (10.16)								
	24.99 (16.11)	32.30 (13.75)	14.49 (8.93)	1.45 (.25)	1.24 (.23)	11.69 (8.28)								

Table 4 Interaction effects of topic, writers' L1 background, and L2 proficiency

Feature	Topic × L1 × Proficiency			Topic × L1			Topic × Proficiency			L1 × Proficiency		
	<i>F</i>	<i>p</i>	η_p^2	<i>F</i>	<i>p</i>	η_p^2	<i>F</i>	<i>p</i>	η_p^2	<i>F</i>	<i>p</i>	η_p^2
Hedge	.67	.616	.003	5.71	.003*	.011	1.20	.302	.002	3.09	.015	.012
Booster	.57	.682	.002	.05	.056	.000	1.90	.150	.004	.34	.849	.001
Attitude	2.16	.072	.009	13.32	< .001*	.026	1.06	.348	.002	.74	.568	.003
Self	.35	.845	.001	18.86	< .001*	.037	1.10	.333	.002	1.65	.159	.007
Reader	2.02	.089	.008	3.58	.028	.007	.97	.381	.002	1.51	.198	.006
Directive	.68	.605	.003	15.53	< .001*	.031	2.24	.107	.005	.50	.738	.002

**p* values are significant with the Bonferroni correction ($p < .05/6$ or .0083)

measure for research including within-subject variables. Cohen (1988) suggested η_p^2 values of .0099, .0588, and .1379 as small, medium, and large effects, respectively. However, given the suggestion that effect sizes should be interpreted differently in different research fields (Ferguson, 2009; Plonsky & Oswald, 2014), the use of effect size values in this study is rather restricted to their relative sizes across different measures.

Results

Table 4 presents the interaction effects of the three independent variables on interactional metadiscourse features obtained from three-way ANOVAs. As shown in the first column, there was no significant three-way interaction among the effects of topic, L1 background, and L2 proficiency on any of the metadiscourse categories. The result also showed no significant two-way interaction effect between topic and L2 proficiency, nor between L1 background and L2 proficiency. However, there were significant interaction effects between topic and L1 background on various metadiscourse categories (i.e., hedges, attitude markers, self-mentions, and directives), which means that patterns of topic effects on EFL students' interactional metadiscourse markers are likely to differ across their L1 backgrounds. The interaction effects of topic and L1 background were mostly small in size, with η_p^2 ranging from .011 to .037.

Post-hoc analyses showed that the Chinese student group used significantly more hedges in the Smoking topic ($p < .001$, $d = -.203$),¹ while the Japanese and Korean groups displayed no significant difference in their hedge use across the topics (Japanese: $p = .995$, $d < .001$; Korean: $p = .050$, $d = .132$). In terms of attitude markers, the Korean group were the only group that showed a significant topic difference (i.e., increased use of attitude markers in the Smoking topic; $p < .001$, $d = -.286$).

¹ Cohen's *d* values indicate effect sizes for the comparison between two means. Cohen (1988) set *d* values of .2, .5, and .8 as benchmarks for small, medium, and large effects.

Table 5 Main effects of topic, writers' L1 background, and L2 proficiency

Feature	Topic			L1			Proficiency		
	<i>F</i>	<i>p</i>	η_p^2	<i>F</i>	<i>p</i>	η_p^2	<i>F</i>	<i>p</i>	η_p^2
Hedge	.20	.654	.000	43.61	< .001*	.081	2.82	.060	.006
Booster	351.24	< .001*	.263	28.59	< .001*	.055	.66	.519	.001
Attitude	3.42	.065	.003	10.68	< .001*	.021	1.41	.245	.003
Self	53.54	< .001*	.052	65.43	< .001*	.117	1.87	.155	.004
Reader	72.03	< .001*	.068	34.16	< .001*	.065	3.94	.020	.008
Directive	17.86	< .001*	.018	20.52	< .001*	.040	.80	.452	.002

**p* values are significant with the Bonferroni correction ($p < .05/6$ or .0083)

Furthermore, the Japanese and Korean groups increased their use of self-mentions in the Part-time topic (Japanese: $p < .001$, $d = .387$; Korean: $p < .001$, $d = .326$), but the Chinese students did not change their use of self-mentions across the topics ($p = .742$, $d = -.017$). The Chinese and Korean groups showed significantly increased use of directives in the Smoking topic (Chinese: $p < .001$, $d = -.266$; Korean: $p < .001$, $d = -.338$), while the Japanese group used significantly more directives in the Part-time topic ($p = .001$, $d = .167$). These results indicating the topic effect dependent on the writer's L1, although statistically significant, should be interpreted cautiously because of the small magnitude of all interaction effects (η_p^2 from .011 to .037), and the main effect analysis in fact showed greater differences in the frequency of metadiscourse markers (medium to large effects).

Table 5 presents the main effects of each independent variable on metadiscourse categories. The result showed statistically significant effects of topic and L1 background on interactional metadiscourse. With regard to topic differences, it was found that EFL students used significantly more boosters and directives in the Smoking topic (boosters: $p < .001$, $\eta_p^2 = .263$; directives: $p < .001$, $\eta_p^2 = .018$), while significantly more self-mentions and reader pronouns in the Part-time topic (self-mentions: $p < .001$, $\eta_p^2 = .052$; reader pronouns: $p < .001$, $\eta_p^2 = .068$), with the largest effect size for boosters. These results suggest that topic may play an important role in directing the writer's attention to varying dimensions of interactional metadiscourse, offering evidence of notable topic effects on interpersonal discourse features, aside from earlier research findings that have suggested topic effects on syntactic and lexical measures. Extracts 1 and 2 show some instances of boosters and directives in the Smoking topic essays.

Extract 1 (KOR-SMK-192-B12)

Smoking **must** be **completely** banned at **all** the restaurants in the country. It's not **should** be banned, it **must** be **completely** banned. As you know, there are **so many** smokers live in our world. And also there are few non-smokers live here. But you see, there is no place for non-smokers.

Table 6 Summary of post hoc comparisons across L1 differences

Feature	L1 differences
Hedge	Japanese > Chinese > Korean
Booster	Japanese = Chinese > Korean
Attitude	Japanese > Korean = Chinese
Self-mention	Japanese > Korean > Chinese
Reader	Chinese > Japanese = Korean
Directive	Japanese > Korean = Chinese

Inequality signs indicate statistical significance with the Bonferroni correction

Extract 2 (CHN-SMK-197-B11)

So it's **necessary** to ban smoking **completely** at **all** the restaurants in the country. In my opinion, smoking in the public, **especially** places like restaurant, is a immoral behavior. It's **really** a bad experience that there has smoker around you when you were eating.

Extract 3 demonstrates some instances of self-mentions in the Part-time topic essays.

Extract 3 (KOR-PTJ-185-B12)

When **I** was a high school student 3 grade, **I** was doing a part time job. Then **I** met many people and **I** feel not good. Because guest is not good, work is very hard and it made **me** very tired. **I** did not know that time is priceless experience.

The result also showed that the EFL students have different patterns of metadiscourse use across their L1 backgrounds. As displayed in Table 5, L1 background effects were statistically significant on all types of metadiscourse categories, with the largest effect on self-mentions ($\eta_p^2 = .117$). Post-hoc analyses with the Bonferroni correction showed that the Japanese group tended to pay more attention to interactional metadiscourse than the Chinese and Korean groups, except only for reader pronouns (see Table 6 for summary). Specifically, the Japanese group was found to use more hedges than the other student groups; of the two remaining groups, the Chinese students used more hedges than the Korean students. The Korean student group also exhibited the fewest number of boosters, while the Japanese and Chinese groups showed comparable usage of boosters. Another notable finding is that, compared to the other groups, the essays by the Chinese students included the fewest number of self-mentions but the highest number of reader pronouns, which might be a reflection of their stronger ideology that emphasizes the importance of communities and groups over individuals (Bloch & Chi, 1995; Shen, 1989). Taken together, it can be suggested that East Asian EFL students with different L1 backgrounds have remarkably different patterns of interactional metadiscourse use, thus highlighting the need to avoid a monolithic conceptualization of East Asian learners of English (e.g., learners from a collectivist culture) in the context of L2 research.

Table 7 Summary of significant differences between the EFL student groups and the NS Group

Topic	L1	Hedge	Booster	Attitude	Self	Reader	Directive
Part-time	Chinese	–			–	+	
	Japanese	–			+	+	+
	Korean	–		–		+	
Smoking	Chinese	–	+		–	+	
	Japanese	–	+	+		+	+
	Korean	–		+		+	

EFL students at the B1.1 level

–Indicates a value significantly lower than the value for the NS group ($p < .05/18$ or $.0028$) and + Indicates a value significantly higher than the NS group value ($p < .05/18$ or $.0028$) with the Bonferroni correction

Unlike the results of topic and L1 background, there was no significant effect of L2 proficiency on interactional metadiscourse. That is, the EFL students at varying proficiency levels showed a similar pattern of interactional features in argumentative writing. This finding can be interpreted in entirely different ways depending on how the ideal use of interactional metadiscourse is set. One of the possible explanations is that East Asian EFL students are likely to develop a good command of interactional metadiscourse at low L2 proficiency, particularly in terms of their extensive use, and accordingly they do not show further changes in the quantity of interactional metadiscourse as their proficiency level advances. An alternative explanation is that East Asian EFL students, even those at an intermediate-high level, have difficulty in noticing important functions of interactional metadiscourse and improving their ability to use interpersonal markers in written discourse. Given these competing interpretations, I further analyzed the essays composed by 200 native English-speaking (NS) college students under similar writing constraints. This dataset, a part of the ICNALE corpus, constitutes 400 essays on the same writing topics (i.e., Part-time and Smoking topics). For this analysis, I conducted a series of independent *t*-tests to reveal significant differences between the NS group and the EFL student groups at the highest proficiency level (B1.2) for each topic, which would enable me to avoid redundant analyses of L1 background or L2 proficiency effects.

As summarized in Table 7, there were several notable distinctions between the NS group and the EFL student groups. The results showed that the NS group used significantly more hedges than all EFL student groups (Chinese: $p < .001$, $d = .96$; Japanese: $p < .001$, $d = .53$; Korean: $p < .001$, $d = .99$ for Part-time; Chinese: $p < .001$, $d = .56$; Japanese: $p = .001$, $d = .36$; Korean: $p < .001$, $d = .86$ for Smoking). Additionally, the NS group was found to use significantly more self-mentions than the Chinese group ($p < .001$, $d = .58$ for Part-time; $p < .001$, $d = .67$ for Smoking). On the other hand, the NS group used a fewer number of reader pronouns than the EFL student groups (Chinese: $p < .001$, $d = -1.42$; Japanese: $p < .001$, $d = -.80$; Korean: $p < .001$, $d = -.68$ for Part-time; Chinese: $p < .001$, $d = -1.11$; Japanese: $p = .001$, $d = -.33$; Korean: $p < .001$, $d = -.52$ for Smoking). It was also found that the NS group used fewer directives than the Japanese group ($p < .001$, $d = -.79$ for

Part-time; $p = .001$, $d = -.65$ for Smoking). These results are further explored from a developmental perspective in the Discussion section.

Discussion

This study explored the effects of one task variable (i.e., topic) and two learner variables (i.e., L1 background and L2 proficiency) on the quantity of interactional metadiscourse in written argumentation. There were significant interaction effects between topic and L1 background on hedges, attitude markers, self-mentions, and directives, indicating that topic effects on the occurrences of these metadiscourse markers differed by East Asian learners' L1. However, as reported above, the size of the interaction effects was generally smaller than that of each of their main effects (topic and L1 background individually), which will be discussed in greater depth. The present results are similar to the findings reported in Hinkel (2009); examining the essays written by Chinese, Japanese, and Korean EFL students, Hinkel revealed that topic tended to affect their use of modal verbs in a consistent manner, but some topics exerted a much greater impact on a particular L1 group (e.g., Japanese learners' overuse of possibility and ability modals in the essays on the selection of academic majors).

With regard to the research question related to topic effects, I found a clear difference in the EFL students' use of interactional metadiscourse across the two topics, confirming the role of topic as one of the factors that elicits different interpersonal discourse features in writing. Specifically, the study found that the Part-time topic allowed L2 writers to express their authorial presence and awareness of imagined readers more explicitly (increased use of self-mentions and reader pronouns), while L2 writers' opinions were more strongly expressed through increased use of boosters and directives for the Smoking topic. These findings may give support to some widely acknowledged patterns of discourse that Asian students tend to reduce explicit authorial presence expressed by personal pronouns when making strong arguments (Scollon & Scollon, 2001). Additionally, given these diverging directions of topic effects, I tentatively argue for the need to understand the construct of interactional metadiscourse in ways that reflect how interactional markers are actually used; some categories of interactional metadiscourse are closely linked (e.g., boosters and directives contributive to an assertive voice; self-mentions and reader pronouns for explicit writer-reader interaction), but such patterns of co-occurrences do not match the existing sides of Hyland's (2005) interactional metadiscourse framework, namely, stance and engagement.

The finding of topic differences can be further discussed with regard to prompt wording. As illustrated in the Results section, boosters were found to be the interactional metadiscourse category with the largest topic effect (see Table 5), and looking into the exact wording of the Smoking topic informs us that its 12-word statement includes as many as three boosters (i.e., Smoking *should be completely* banned at *all* the restaurants in the country). Given this excessive inclusion of boosters in the Smoking topic, it is plausible that the EFL students might have been primed to use more boosters when responding to this topic. This issue can be of great

importance in the context of testing because the inclusion of some stance markers in a prompt can implicitly lead test-takers to compose their essays with a particular point of view, potentially affecting their language and performance scores negatively (Hudson, Lane, & Mercer, 2005; Korbin, Deng, & Shaw, 2011). Therefore, writing prompts to be used in high-stakes test settings should be constructed with few biased or emotional words, unless the elicitation of particular language features is intentionally planned, like facilitating the use of hedging expressions.

My argument on priming effects of prompt wording on interactional metadiscourse is still speculative because this study did not make a direct comparison between writing samples from a prompt that includes, for example, multiple boosters with those from the same prompt without boosters. With further research with systematically manipulated prompts, we will be able to confirm if the manipulation of prompt wording functions as an effective strategy for L2 writing development. For example, when writing instruction aims to promote students' production of hedges, students can be provided with short writing tasks that do and do not include several hedges in their prompts. Students who have experienced various writing tasks with and without hedges would raise awareness of the roles of hedges and, in turn, learn how to produce them appropriately. This implicit process of acquiring how to use hedging expressions can be followed by a more explicit post-writing phase that allows students to discuss various functions of hedges in making arguments.

Additionally, this study demonstrated clear differences in the use of interactional features across the different L1 groups. The essays written by the Korean students were characterized by fewer occurrences of hedges and boosters, while the essays by the Japanese students overall included greater numbers of stance and engagement markers than those by the other groups. Some notable features of the Chinese student essays, compared with those by the other groups, were a greater number of reader pronouns and a fewer number of self-mentions. These differences between the groups may be associated with how they build their arguments in persuasive discourse (Williams, 2017). Japanese speakers are likely to use more explicit linguistic markers that entail the functions of potentiality and ambiguity (Maynard, 1993), constructing their stance toward a proposition using an increased number of metadiscourse markers as a means of being polite and opening alternative ideas (Holmes, 1982). On the other hand, in an effort to avoid ambiguity or uncertainty (Paek, 2005), Korean speakers tend to focus on conveying their ideas and arguments straightforwardly with reduced attention to epistemic stance or engagement with readers. These findings may point to the importance of going beyond the broad categorization of East Asian learners of English as those who have a similar cultural and linguistic background in L2 pedagogy and research. For example, Shen (1989) found that Chinese students experience difficulties with first person pronouns in their English essays. Hong and Cao (2014) suggested that Chinese EFL students tend to use more interpersonal discourse markers in writing than Spanish and Polish EFL students. Given the large proportion of Chinese speakers to the entire English learner population, these findings have often been extrapolated to other East Asian learners of English. However, as found in this study, East Asian EFL students with different L1 backgrounds show notably distinct patterns of interactional metadiscourse features in their writing.

Contrary to the findings related to topic and L1 background, this study showed that the learner essays did not contain different numbers of interactional meta-discourse features depending on L2 proficiency. This finding, together with the results of clear L1 differences, provides support to what Jarvis (2000) suggested in his research on L1 influence. Given the finding of a remarkable similarity in vocabulary use among learners with the same L1, Jarvis argued “learners from the same L1 background, despite differences in age and L2 exposure, exhibit higher levels of homogeneity than learners from different L1 backgrounds who are matched according to age and L2 exposure” (p. 289). Confirming the important role of L1 background in L2 use, this study may also offer evidence that the essays written by the same-L1 learners, regardless of their L2 proficiency, are likely to display similar patterns of stance and engagement markers.

Comparing the L2 learner results with those from the NS group, this study demonstrated that L2 learner writing is likely to contain fewer hedges but more reader pronouns than NS writing, potentially identifying what categories of interactional metadiscourse we should attend to as a pedagogical target. I acknowledge that it is not ideal to use the essays composed by NS college students as the norm because NS students’ writing proficiency would also vary depending on their experience with academic writing (Hall & Navarro, 2011; Römer, 2009). Nevertheless, the NS essays used in this study might have served as the benchmark data more effectively than other potential datasets in that they were based on the same topics as those used for the EFL essays and that the L1 English-speaking college students had gained more *English* academic writing experience than the EFL students. Furthermore, previous studies indicated that the realistic aim of many L2 writers is to develop the ability to write essays similar to those of their NS counterparts (e.g., Zhou, Busch, & Cumming, 2014), justifying the methodological decision of using the NS student essays as the norm.

This study, similar to earlier research (e.g., Hyland & Milton, 1997; Lee & Deakin, 2016), revealed that L2 learners tend to underuse hedging markers in writing as compared to their NS counterparts, potentially leading to more straightforward, stronger assertions. In this study, the reduced use of hedges was identified as a common feature of all L2 learner groups, pointing to the need to include hedging as a target trait in L2 writing instruction. Additionally, the L2 groups’ greater use of reader pronouns than the NS group can be interpreted as the outcome of their cultural backgrounds that value harmony with other members in a community and regard the obvious projection of the self as being rude or disrespectful (Bloch & Chi, 1995; Shen, 1989). However, it should be noted that, while writers’ use of reader pronouns plays important roles in establishing a common ground with readers (Hyland, 2001), their overreliance on personal pronouns can still cause harm to the formality of academic writing. When observing any overuse of reader pronouns in L2 learner writing, writing instructors may need to make an effort to have their students notice its potential drawbacks by, for example, providing them with the opportunity to compare their essays containing an excessive number of reader pronouns with model essays (e.g., those reformulated by the instructor) that contain an appropriate number of reader pronouns.

This study offers insight into how to design a lesson on interactional metadiscourse, with a clear understanding of varying stance and engagement styles of different L1 groups. Based on the current findings, for example, a course for Korean students can focus on encouraging them to build more convincing and balanced arguments with increased hedge and booster markers; a course for Chinese students can be aimed at having students develop their ideas using more self-mentions so that they express their individual voices in writing more clearly. Finally, given the finding of no significant difference between learners at different levels, it can be suggested that (1) we should attend to the diversity or appropriacy of interactional metadiscourse markers rather than their quantity to capture development, or (2) interactional metadiscourse is a multidimensional construct whose development is not commensurate with the general improvement of L2 proficiency, possibly due to its optionality in context (i.e., an area with no binary outcome of being correct or incorrect). Nevertheless, as evidenced by earlier studies (e.g., Abbuhl, 2006; Cheng & Steffensen, 1996), the ability to use interactional metadiscourse markers can be promoted through focused instruction, indicating the need to explore the effectiveness of explicit and interventionist instruction for L2 learners' extensive, diverse, and appropriate use of interactional metadiscourse features.

Conclusion

This study showed that L2 learners' use of interactional metadiscourse differed across topics and L1 backgrounds. Based on the result of topic differences, the study discussed that when designing and/or selecting writing prompts, teachers and testing experts should take into account potential effects of prompt orientation and prompt wording on interactional metadiscourse. The findings related to learner variables provided valuable insight into what the development of interactional metadiscourse actually means and how to foster it in instructional contexts. Nevertheless, there are several limitations of this study that are worth noting for the careful interpretation of its findings. First, this study approximated proficiency improvement through a cross-sectional analysis of the corpus data. We may need a follow-up longitudinal study on the development of interactional metadiscourse awareness and use in writing to trace actual developmental trajectories over time. Also, this study examined the development, or significant changes, of interactional metadiscourse from a quantitative approach that focused primarily on the number of interactional metadiscourse markers in each text. This quantity-based analysis may fail to provide a full picture of learners' interactional metadiscourse competence, so it would be of great importance to identify what types of interactional markers (within each metadiscourse category) are prevalent in learner writing and how the use of such frequently used markers varies with L2 proficiency. While the discussion of topic effects on interactional metadiscourse provides insight into teaching and assessment, they should be accepted with caution because this, as a large-scale corpus study, was based only on the two topics and there was no systematic manipulation of topic differences. Also, essays composed under time constraints may not fully reflect learners' true writing ability because timed writing rarely involves revising or editing. Future research

with many carefully modified prompts and extended time will be able to address these limitations and yield more generalizable findings. Interactional metadiscourse has become an important area in L2 writing and EAP research, as the writer's proper interaction with the audience has been identified as one of the keys to successful academic writing. Future research addressing the points discussed above will enable us to achieve a greater understanding of how to construct an effective authorial voice in a variety of academic settings.

Compliance with ethical standards

Conflict of interest The author declares that they have no conflict of interest.

References

- Abbuhl, R. (2006). Hedging and boosting in advanced-level L2 legal writing: The effect of instruction and feedback. In H. Byrnes, H. Weger-Guntharp, & K. A. Sprang (Eds.), *Educating for advanced foreign language capacities: Constructs, curriculum, instruction, assessment* (pp. 152–164). Washington, DC: Georgetown University Press.
- Aull, L. (2019). Linguistic markers of stance and genre in upper-level student writing. *Written Communication, 36*, 267–295.
- Aull, L. L., & Lancaster, Z. (2014). Linguistic markers of stance in early and advanced academic writing: A corpus-based comparison. *Written Communication, 31*, 151–183.
- Bax, S., Nakatsuhara, F., & Waller, D. (2019). Researching L2 writers' use of metadiscourse markers at intermediate and advanced levels. *System, 83*, 79–95.
- Bloch, J., & Chi, L. (1995). A comparison of the use of citations in Chinese and English academic discourse. In D. Belcher & G. Braine (Eds.), *Academic writing in a second language: Essays on research and pedagogy* (pp. 231–274). Norwood: Ablex.
- Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing, 26*, 42–65.
- Çandarlı, D., Bayyurt, Y., & Martı, L. (2015). Authorial presence in L1 and L2 novice academic writing: Cross-linguistic and cross-cultural perspectives. *Journal of English for Academic Purposes, 20*, 192–202.
- Cheng, X., & Steffensen, M. S. (1996). Metadiscourse: A technique for improving student writing. *Research in the Teaching of English, 30*, 149–181.
- Cheung, Y. L., & Low, T. H. (2019). Pre-university students' voice construction in argumentative essays. *RELC Journal, 50*, 269–284.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale: Lawrence Erlbaum.
- Council of Europe. (2018). *Common European framework of reference for languages: Learning, teaching and assessment*. Cambridge: Cambridge University Press.
- Crossley, S. A., & McNamara, D. S. (2014). Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners. *Journal of Second Language Writing, 26*, 66–79.
- Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice, 40*, 532–538.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). London: SAGE.
- Hall, J., & Navarro, N. (2011). Lessons for WAC/WID from language learning research: Multicompetence, register acquisition, and the college writing student. *Across the Disciplines, 8*, 1–16.
- Hinds, J. (1990). Inductive, deductive, quasi-inductive: Expository writing in Japanese, Korean, Chinese, and Thai. In U. Connor & A. M. Johns (Eds.), *Coherence in Writing: Research and pedagogical perspectives* (pp. 89–109). Alexandria: Teachers to Speakers of Other Languages.

- Hinkel, E. (2003). Adverbial markers and tone in L1 and L2 students' writing. *Journal of Pragmatics*, 35, 1049–1068.
- Hinkel, E. (2009). The effects of essay topics on modal verb uses in L1 and L2 academic writing. *Journal of Pragmatics*, 41, 667–683.
- 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.
- Holmes, J. (1982). Expressing doubt and certainty in English. *RELC Journal*, 13, 9–28.
- Hong, H., & Cao, F. (2014). Interactional metadiscourse in young EFL learner writing: A corpus-based study. *Interactional Journal of Corpus Linguistics*, 19, 201–224.
- Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the 10th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 168–177).
- Hudson, R. F., Lane, H. B., & Mercer, C. D. (2005). Writing prompts: The role of various priming conditions on the compositional fluency of developing writers. *Reading and Writing: An Interdisciplinary Journal*, 18, 473–495.
- Huot, B. (1990). The literature of direct writing assessment: Major concerns and prevailing trends. *Review of Educational Research*, 60, 237–263.
- Hyland, K. (2001). Bringing in the reader: Addressee features in academic articles. *Written Communication*, 18, 549–574.
- Hyland, K. (2002a). Directives: Argument and engagement in academic writing. *Applied Linguistics*, 23, 215–239.
- Hyland, K. (2002b). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics*, 34, 1091–1112.
- Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, 7, 173–192.
- Hyland, K., & Milton, J. (1997). Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language Writing*, 6, 183–205.
- Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied Linguistics*, 25, 156–177.
- Institute of International Education (IIE). (2019). *Open doors 2019 fast facts*. Retrieved from <https://www.iie.org/Research-and-Insights/Open-Doors/Fact-Sheets-and-Infographics/Fast-Facts>.
- Intaraprawat, P., & Steffensen, M. S. (1995). The use of metadiscourse in good and poor ESL essays. *Journal of Second Language Writing*, 4, 253–272.
- Ishikawa, S. (2013). The ICNALE and sophisticated contrastive interlanguage analysis of Asian learners of English. In S. Ishikawa (Ed.), *Learner corpus studies in Asia and the world* (pp. 91–118). Kobe: Kobe University.
- Jarvis, S. (2000). Methodological rigor in the study of transfer: Identifying L1 influence in the interlanguage lexicon. *Language Learning*, 50, 245–309.
- Klein, D., & Manning, C. D. (2003). Accurate unlexicalized parsing. In *Proceedings of the 41st meeting of the association for computational linguistics* (pp. 423–430).
- Korbin, J. L., Deng, H., & Shaw, E. J. (2011). The association between SAT prompt characteristics, response features, and essay scores. *Assessing Writing*, 16, 154–169.
- Lee, J. J., & Casal, J. E. (2014). Metadiscourse in results and discussion chapters: A cross-linguistic analysis of English and Spanish thesis writers in engineering. *System*, 46, 39–54.
- Lee, J. J., & Deakin, L. (2016). Interactions in L1 and L2 undergraduate student writing: International metadiscourse in successful and less-successful argumentative essays. *Journal of Second Language Writing*, 33, 21–34.
- Li, T., & Warton, S. (2012). Metadiscourse repertoire of L1 Mandarin undergraduates writing in English: A cross-contextual, cross-disciplinary study. *Journal of English for Academic Purposes*, 11, 345–356.
- Maynard, S. (1993). *Discourse modality: Subjectivity, emotion and voice in the Japanese language*. Amsterdam: John Benjamins.
- Mohammad, S., & Turney, P. (2013). Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29, 436–465.
- Paek, H. (2005). Understanding celebrity endorsers in cross-cultural contexts: A content analysis of South Korean and US newspaper advertising. *Asian Journal of Communication*, 15, 133–153.
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64, 878–912.

- Qin, J., & Karabacak, E. (2010). The analysis of Toulmin elements in Chinese EFL university argumentative writing. *System*, 38, 444–456.
- Reid, J. (1990). Responding to different topic types: A quantitative analysis from a contrastive rhetoric perspective. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 191–210). New York: Cambridge University Press.
- Römer, U. (2009). English in academia: Does nativeness matter? *Anglistik: International Journal of English Studies*, 20, 89–100.
- Scollon, R., & Scollon, S. W. (2001). *Intercultural communication* (3rd ed.). Oxford: Blackwell.
- Shen, F. (1989). The classroom and the wider culture: Identity as key to learning English composition. *College Composition and Communication*, 40, 456–466.
- Williams, C. H. (2017). *Teaching English in East Asia: A teacher's guide to Chinese, Japanese, and Korean learners*. Singapore: Springer.
- Wolfe, C. R. (2011). Argumentation across the curriculum. *Written Communication*, 28, 193–219.
- Yang, W., & Kim, Y. (2020). The effect of topic familiarity on the complexity, accuracy, and fluency of second language writing. *Applied Linguistics Review*, 11, 79–108.
- Yang, W., Lu, X., & Weigle, S. C. (2015). Different topics, different discourse: Relationships among writing topic, measures of syntactic complexity, and judgments of writing quality. *Journal of Second Language Writing*, 28, 53–67.
- Yoon, H. (2017a). Textual voice elements and voice strength in EFL argumentative writing. *Assessing Writing*, 32, 72–84.
- Yoon, H. (2017b). Linguistic complexity in L2 writing revisited: Issues of topic, proficiency, and construct multidimensionality. *System*, 66, 130–141.
- Yoon, H., & Römer, U. (2020). Quantifying disciplinary voices: An automated approach to interactional metadiscourse in successful student writing. *Written Communication*, 37, 208–244.
- Zhao, C. G. (2017). Voice in timed L2 argumentative essay writing. *Assessing Writing*, 31, 73–83.
- Zhou, A. A., Busch, M., & Cumming, A. (2014). Do adult ESL learners' and their teachers' goals for improving grammar in writing correspond? *Language Awareness*, 23, 234–254.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.