LIF – Language in Focus Journal, Volume 2, Issue 1, 2016, DOI: 10.1515/lifijsal-2016-0002

Investigating Metadiscourse Markers in Asian Englishes: A Corpus-Based Approach

Yuichiro Kobayashi
Toyo University, Japan

Abstract

The present study investigated differences in rhetorical preferences in L2 writings among different L1 groups. This study compared the use of metadiscourse markers in L2 essays and identified discourse devices used to distinguish different L1 groups. The essays originated from the International Corpus Network of Asian Learners of English (ICNALE) compared six L1 groups (viz., Chinese, Indonesian, Japanese, Korean, Taiwanese, and Thai) based on the frequency of metadiscourse markers. I utilized heat map with hierarchical clustering to investigate differences in metadiscourse among the six learner groups. The results suggested a substantial difference in the use of metadiscourse markers between East Asian groups (viz., Chinese, Japanese, Korean, and Taiwanese) and Southeast Asian groups (viz., Indonesian and Thai). Furthermore, each learner group displayed the specific characteristics of metadiscourse, which offer suggestions for improving L2 learners' writings.

1. Introduction

As globalization has increased intercultural and interlingual contacts, it is increasingly important to understand the varieties of English as foreign languages. Cultural differences in language have been the main topic of contrastive rhetoric, which identifies the writer's first language (L1) transfer to second language (L2) writing in terms of rhetorical strategy (Conner, 1996). Rhetorical preferences in L1 can affect various aspects of L2, such as paragraph development (Bickner & Peyasantiwong, 1988), discourse development (Reid, 1992), and metadiscourse (Crismore, Markkanen, & Steffensen, 1993).

Among these aspects, metadiscourse attracts the most attention in current linguistic research and language teaching.

Corpus-based studies on academic writings highlight the importance of metadiscourse to improving written communication (Dahl, 2004; Hyland & Tse, 2004; Kuhi & Behnam, 2011). By using computerized learner corpora, linguists can obtain a large amount of frequency-based information on metadiscourse, which reveals overuse and underuse patterns across the interlanguages of different L1 groups. Accordingly, it can be utilized to suggest whether or not L2 metadiscourse is affected by the L1. Despite the diversity of interlanguages, academic writing has a set of preferred rhetorical conventions. Therefore, learners must conform to the conventions and acquire appropriate discourse styles.

2. Literature Review

2.1 Contrastive Interlanguage Analysis

Since the development of learner corpus research in the late 1990s, contrastive interlanguage analysis has grown rapidly as a leading method in the field. It consists of two types of comparison: (a) comparison of native language and interlanguage and (b) comparison of different interlanguages (Granger, 1996). Numerous studies on contrastive interlanguage analysis employed the International Corpus of Learner English (ICLE), which contains 3.7 million words of writing samples from 16 native language backgrounds. The corpus has a comparable data set, the Louvain Corpus of Native English Essays (LOCNESS), which contains 324 thousand words of native writers' essays. Using these two corpora, Granger and Rayson (1998) compared the use of nine word categories in essays written by French learners and native speakers, and showed that French learners used a number of linguistic features characteristic of spoken language. Aijmer (2002) also examined the frequencies of modal devices in native speakers' and Swedish learners' writings, and revealed the learners' overuse of all the modal categories examined in her paper.

Researchers conducted extensive contrastive interlanguage studies throughout the world with a focus on the use of vocabulary and certain grammatical features. Consequently, there are few corpus-based discourse analyses due to the difficulty of dealing with language characteristics that extend across clause boundaries (Biber, Conrad, & Reppen, 1998). However, a particular list of discourse items, such as logical connectors or stance markers, enables researchers to conduct corpus-based discourse studies (Conrad, 2002). For instance, Altenberg and Tapper (1998) utilized a list of adverbial connectors to

analyze the logical structure in argumentative essays written by Swedish and French learners of English. Biber (2006) also applied a list of common lexico-grammatical features, particularly those associated with writer's stance, to investigate spoken and written activities in academic life. These studies indicate the potential of a corpus-based approach to L2 discourse.

2.2 Metadiscourse Analysis

By applying the methodology of contrastive interlanguage analysis to a wide variety of learner corpus research, linguists identified several linguistic features characteristic of different learners groups. A relatively new area of investigation in contrastive interlanguage analysis is the study of metadiscourse. For example, Ädel (2006) revealed the overuse of both personal and impersonal metadiscourse markers in Swedish learners' English. Hong and Cao (2014) also shed light on the differences in the use of interactional metadiscourse markers among Chinese, Polish, and Spanish learners of English. Moreover, Tan and Eng (2014) and Lin (2014) investigated the use of metadiscourse markers in the writing of Malaysian English learners and in the speech of Chinese English learners, respectively. The results of these studies suggest that learners' L2 performance is influenced by cultural factors prevalent in their L1 communities. However, few studies have examined the metadiscourse of multiple learner groups from different language backgrounds. Rather, most of the previous studies compared native language with a particular interlanguage. Thus, it is necessary to contrast different learner groups to gain a deeper insight into L1-induced rhetorical differences in L2 performance. By investigating multiple learner groups, researchers can determine whether certain characteristics of metadiscourse are universal phenomena or unique traits indigenous to a specific L1.

3. Research Design

3.1 Purpose of the Study

The present study aimed to investigate differences in rhetorical preferences in L2 writings among different L1 groups. This study compares the use of metadiscourse markers in L2 essays and identifies discourse devices that can be used to distinguish different L1 groups. The findings of this study can provide a pedagogical suggestion for effective writing instruction for each learner group.

3.2 Corpus Data

The present study draws on the written component of the International Corpus Network of Asian Learners of English (ICNALE-Written) which contains 1.3 million words of argumentative essays written by 2,600 college students in ten Asian countries and areas (Ishikawa, 2013). The data analyzed in this study is a subset from this corpus, including the written compositions of six L1 groups (viz., Chinese, Indonesian, Japanese, Korean, Taiwanese, and Thai). In the viewpoint of World Englishes, these groups consist of an expanding circle of English users. The subset includes only writers with a B1 CEFR level. The writing conditions and learners' proficiency levels were strictly controlled for the comparison of these groups. All essays in the subset were written in response to a single prompt, namely "It is important for college students to have a part-time job" (Ishikawa, 2013, p. 97). Table 1 shows the size of the six learner groups compared in this study.

Table 1. Corpus Size of Six Learner Groups

	Participants	Total words
China (CHN)	337	83,980
Indonesia (IDN)	165	39,096
Japanese (JPN)	228	51,780
Korea (KOR)	149	34,175
Taiwan (TWN)	148	35,294
Thailand (THA)	279	64,186

The ICNALE includes native speakers' essays as well as learners' essays. However, native speakers and learners were not compared in the present study, given the criticism from the theoretical perspective of comparative fallacy (Bley-Vroman, 1983).

3.3 Metadiscourse Markers

The concept of metadiscourse is a matter of controversy amongst linguists, and is considered "a fuzzy term" characterized as "discourse about discourse" or "talk about talk" (Hyland, 2005, p. 16). However, the common thread in definitions of metadiscourse distinguishes propositional and metadiscourse meanings. For instance, according to Williams (1981), metadiscourse is "whatever does not refer to the subject matter being addressed" (p. 226). Crismore, Markkanen, and Steffensen (1993) also defined it as

"[l]inguistic material in texts, written or spoken, which does not add anything to the propositional content but this is intended to help the listener or reader organize, interpret and evaluate the information given" (p. 40). Since metadiscourse is so open to interpretation, it is worth further investigating for its linguistic value.

The present study used the framework of metadiscourse developed by Ken Hyland, which is the most widely accepted in the field of discourse analysis. Hyland (2005) defined metadiscourse as "the cover term for the self-reflective expressions used to negotiate interactional meanings in a text, assisting the writer (or speaker) to express a viewpoint and engage with readers as members of a particular community" (p. 37). Hyland's list of metadiscourse markers is used to analyze different types of texts, such as company annual reports (Hyland, 1998), introductory academic course books (Hyland, 1999), undergraduate textbooks (Hyland, 2000), postgraduate dissertations (Hyland, 2004), and learner writings (Hyland & Tse, 2004). In the present study, I compared six learner groups in terms of the frequency of nearly 500 types of metadiscourse markers listed in Hyland (2005). These metadiscourse resources can be classified into ten functional categories shown in Table 2.

Table 2. Hyland's Classification of Metadiscourse Markers

Category	Function	Examples
Interactive resources	Help to guide reader through the text	
Transitions (TRA)	Express semantic relation between main clauses	*in addition, but, thus, and*
Frame markers (FRM)	Refer to discourse acts, sequences, or text stages	*finally, to conclude, my purpose here is to*
Endophoric markers (END)	Refer to information in other parts of the text	*noted above, see Fig, in section 2*
Evidentials (EVI)	Refer to source of information from other texts	*according to X, (Y, 1990), Z states*
Code glosses (COD)	Help readers grasp functions of ideational material	*namely, e.g., such as, in other words*
Interactional resources	Involve the reader in the argument	
Hedges (HED)	Without writer's full commitment to proposition	*might, perhaps, possible, about*
Boosters (BOO)	Emphasize force or writer's	*in fact, definitely, it is clear*

	certainty in proposition	*that*
Attitude markers (ATM)	Express writer's attitude to proposition	*unfortunately, I agree, surprisingly*
Engagement markers (ENG)	Explicitly refer to or build relationship with reader	*consider, note that, you can see that*
Self-mentions (SEM)	Explicit reference to author(s)	*I, we, my, our*

(Hyland & Tse, 2004, p. 169)

3.4 Statistical Methods

Multivariate statistical methods are very useful when analyzing the relationship between a number of linguistic features and learner groups. These methods are a more statistically sophisticated way to explore the relationship than simple statistical tests, such as the chi-square test or log-likelihood ratio test. Further, multivariate statistical methods can reveal the complex interrelationships among linguistic features and learner groups, along with the association patterns between linguistic features and learner groups. In the field of corpus linguistics, the most common multivariate method is correspondence analysis, which graphically represents the interrelationships amongst linguistic features and learner groups on a two- or three-dimensional scatter plot. However, the scatter plot is often arbitrarily interpreted by manually clustering the linguistic features and learner groups on the plot, and thus, a more reliable clustering technique might be required to better understand the plot.

In the present study, I used heat map with hierarchical clustering to investigate differences in metadiscourse among different L1 groups. This combination of hierarchical cluster analysis and heat map analysis is a powerful method for visualizing multivariate data, such as large frequency tables for corpus-based linguistic analysis (Kobayashi, in press). Furthermore, it can display the results from clustering linguistic features and learner groups, while concurrently generating a heat map from the original frequency table in two-dimensional space. The graphical representation of cluster analysis is easier to interpret than that of correspondence analysis (Glynn, 2014), and the interpretation of the clustering results can be validated with the use of heat map.

3.5 Procedures

The present study calculated the frequencies of ten functional categories of

metadiscourse markers in L2 writings from the six learner groups. Following the frequency counts, this study quantitatively compared the frequencies using heat map with hierarchical clustering, and qualitatively examined the usage examples of metadiscourse markers characteristic of each learner group.

4. Results and Discussion

4.1 Quantitative Analysis

The present study began by calculating the frequencies of ten metadiscourse categories in the writings of six learner groups. Table 3 lists the relative frequencies of categories used by each writer group. The relative frequencies were computed by dividing each cell number by the column total. Since learners infrequently used the endophoric markers, all cell numbers in this row are approximated to zero.

Table 3. The Relative Frequencies of Ten Functional Categories of Metadiscourse Markers

	CHN	IDN	JPN	KOR	TWN	THA
TRA	0.262	0.323	0.235	0.291	0.259	0.276
FRM	0.062	0.054	0.066	0.066	0.058	0.051
END	0.000	0.000	0.000	0.000	0.000	0.000
EVI	0.001	0.001	0.000	0.002	0.001	0.001
COD	0.032	0.064	0.047	0.035	0.048	0.059
HED	0.062	0.054	0.042	0.064	0.072	0.066
BOO	0.098	0.091	0.111	0.099	0.102	0.097
ATM	0.049	0.060	0.071	0.053	0.062	0.050
ENG	0.167	0.165	0.114	0.163	0.173	0.249
SEM	0.267	0.187	0.315	0.227	0.224	0.153

The next step was to investigate the relationships between metadiscourse categories and learner groups through heat map with hierarchical clustering. As shown in Figure 1, the method displayed the results of the clustering of learner groups and metadiscourse categories, as well as generated the heat map from the permutated frequency table in two-dimensional space at the same time. The complete linkage method and Euclidean distances were used to cluster the metadiscourse categories and learner groups.

The results were visualized as tree-like categorizations where small groups of highly similar items are included within much larger groups of less similar items (Oakes, 1998). In the heat map, a comparison is drawn between learner groups, with more frequent metadiscourse categories represented by darker cells, and less frequent categories denoted by lighter cells. Moreover, relative frequencies were placed within each cell in the heat map. By examining the relative frequencies as well as the cell colors in the heat map, the clustering results can be better interpreted in an objective manner.

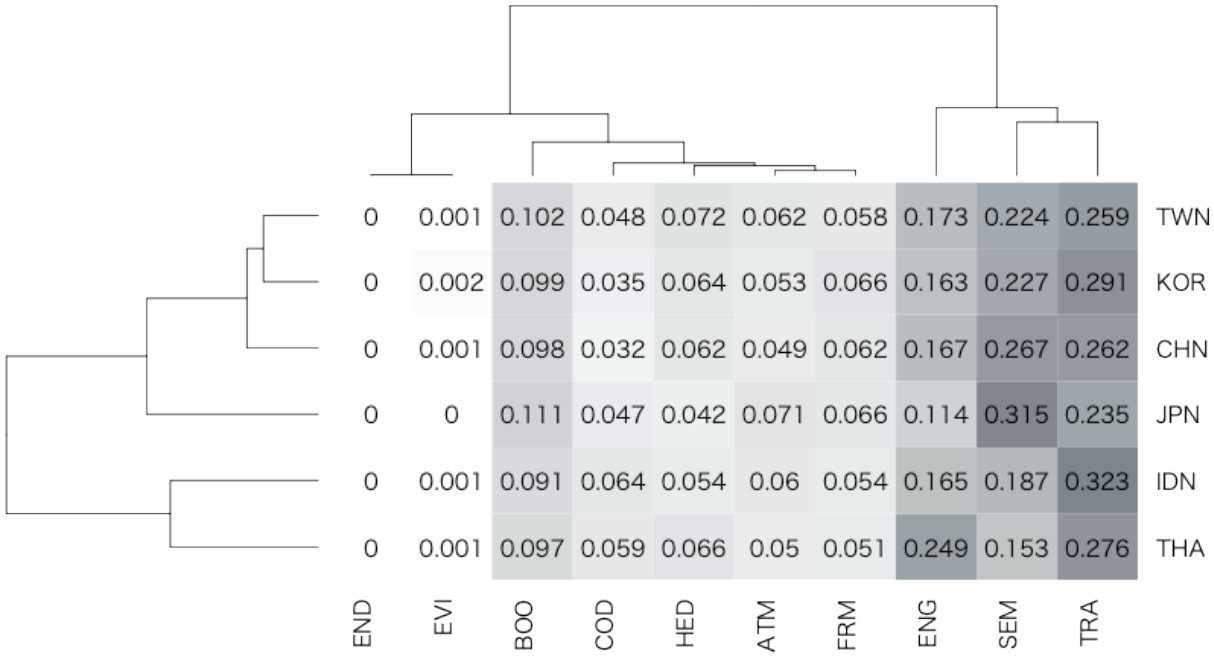

Figure 1. The Relationships between Learner Groups and Metadiscourse Categories

The clustering result of learner groups indicates that there is a substantial difference in the frequency patterns of metadiscourse markers between East Asian groups (viz., Chinese, Japanese, Korean, and Taiwanese) and Southeast Asian groups (viz., Indonesian and Thai). The result of metadiscourse categories showed high-frequency, middle-frequency, and low-frequency categories cluster together respectively. Figure 2 shows the differences of frequencies of metadiscourse categories among learner groups more clearly. In this figure, all frequency information is standardized using *z* scores, which indicates the number of standard deviations each frequency value deviates from the mean of the data set. Therefore, positive and negative scores represent frequencies greater and less than the mean respectively. The lengths of each bar are proportional to the degree of

deviation from the mean. The characteristics of metadiscourse of each learner group will be qualitatively examined in the following sections.

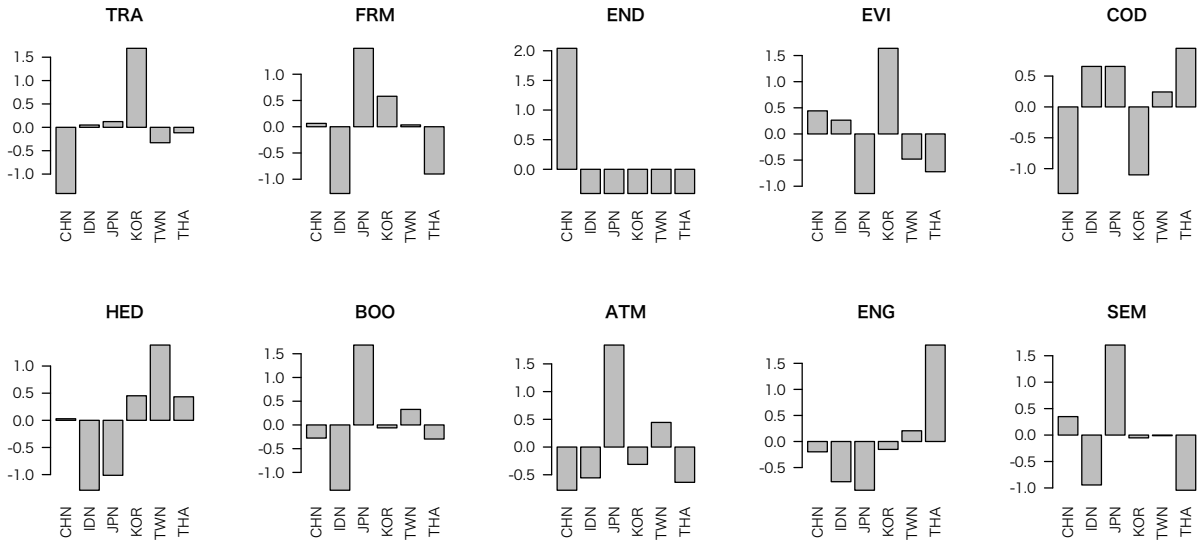

Figure 2. Differences of Metadiscourse Categories among Learner Groups

Notes. All frequency information is standardized using *z* scores.

4.2 Qualitative Analysis

4.2.1 Japanese

As Figure 2 indicates, Japanese learners more frequently used self-mentions, boosters, frame markers, and attitude markers than other groups. Whereas L2 writers whose L1s are reader-responsible languages apt to underuse metadiscourse markers (Kim & Lim, 2013), Japanese learners, who have reader-responsible language background, exhibited the overuse of the above four categories. The most salient feature of Japanese learners was self-mentions, especially *I*, *my*, and *me*. Interestingly enough, they frequently used first person pronouns in their L1 as well as in English, although subject words can be grammatically omitted in Japanese language (Kobayashi, 2011). As Biber, Johanson, Leech, Conrad, and Finegan (1999) illustrate, first person pronouns are linguistic features that characterize spoken language.

(1) **I** am working at a convenience store near **my** home now. (JPN)

(2) The experience was very important for **me**. (JPN)

Another notable feature of Japanese learners was boosters. However, they often used the collocation *I think* as a "softening device" (i.e., hedge) in contrast to native speakers' use of an "emphatic device" (i.e., booster) (Kamimura & Oi, 1998). The fact that different learner groups use identical discourse devices for different argumentative strategies proved very interesting. From a different perspective, spoken language has greatly influenced the heavy use of the words *I think* (Aijmer, 2002).

(3) I **think** that college students should have a part time job. (JPN)
(4) I **think** we should not depend on our parents. (JPN)

In addition, Japanese learners overused frame markers. The overuse may be a consequence of "superficial attention" (Intaraprawat & Steffensen, 1995, p. 271) to logical forms, and result in "artificial, mechanical prose" (Zamel, 1983, p. 27). A frequent use of connectors does not necessarily improve the cohesive quality of a text because semantic relations between main clauses do not have to be marked explicitly (Altenberg & Tapper, 1998).

(5) **First**, college students have much time than junior high school students and working people. (JPN)
(6) **Second**, we have to reduce my parent's expenses. (JPN)

Although there was a notable overuse of attitude markers in Japanese learners' essays, this can be attributed to the wording of the essay prompt itself: "Do you agree or disagree with the following statement? […] It is important for college students to have a part-time job" (Ishikawa, 2013, p. 97). Since the wording of prompt often has a great effect on L2 writings, the reuse of vocabulary and syntactic constructions should be checked meticulously.

(7) I **agree** with the statement. (JPN)
(8) I think that it is **important** for college students to have a part-time job. (JPN)

4.2.2 Korean

Since Korean learners often use coordinate conjunctions at the beginning of sentences, they used a significant number of transitions in their writing. As Biber et al. (1999) demonstrate, *and* and *but* in the sentence-initial position are characteristic of conversation. The use of these conjunctions in academic prose can be attributable to the lack of register awareness as well as the L1 transfer. It is difficult for English learners from Korea and some other Asian countries to distinguish coordination from subordination owing to the use of logical connectors in their L1s (Hinkel, 2002).

(9) **And** parents support their children excessively. (KOR)
(10) **But** I do not think so. (KOR)

Another noteworthy feature of Korean learners was the use of evidentials such as quoting data from surveys or newspapers to justify their arguments in the social context. The prepositional phrase *according to* functions as a hedging device by distancing the writer from the proposition (Biber et al., 1999).

(11) **According to** one survey, 73% of the students are planning to work at a part-time job in this summer vacation. (KOR)
(12) **According to** a newspaper article college student's work part-time wages are low. (KOR)

4.2.3 Chinese

Chinese learners referred to other parts of the text using endophoric markers. In particular, they incorporated supporting details for their argument using *above* and *below* in their persuasive text.

(13) For all these reasons mentioned **above**, it is important for college students to have a part-time job. (CHN)
(14) I have the following reasons **below**. (CHN)

Contrastive rhetorical studies have shown that Chinese writers tend to support their contentions with quotations from classical literature rather than their own ideas (Hinkel, 2002). However, contrary to this rhetorical tradition, Chinese learners analyzed in the

present study are familiar with the manner of English academic writing.

4.2.4 Taiwanese

The prominent feature of Taiwanese learners was hedges, which are one of the most significant rhetorical devices in academic writing. Good writers can use hedges to strengthen their argument and weaken the claim in their discourse (Meyer, 1997).

(15) Some sort of job **would** not be a helpful working experience. (TWN)
(16) I **guess** that many college students have part-time jobs because they need money or more money. (TWN)

Taiwanese learners' use of hedges is possibly influenced by Chinese rhetorical tradition. In Chinese-speaking countries, hedging devices play an important role in arguments because they can concurrently perform several discourse functions due to its ambiguous nature (Hinkel, 2002). In particular, Taiwanese learners purposefully project politeness by using these devices.

4.2.5 Thai

Thai learners used engagement markers, especially second person pronouns, significantly more than other groups. In terms of the writer/reader visibility, they are in contrast to Japanese learners who overused first person pronouns. Engagement markers involve readers as discourse participants whereas self-mentions foreground the presence of author in the text.

(17) Do **you** think this is a good idea? (THA)
(18) **You** can help **your** parents to save their expenditure and **you** can save money for yourself. (THA)

According to Petch-Tyson (1998), second person pronouns are linguistic features that are distinctive of a more informal writing style. In addition, Thai learners frequently used these pronouns in combination with interrogatives, which is also typical of spoken language.

Another characteristic of Thai learners was code glosses. They provided supporting examples using phrases like *such as* and *for example*. Since these expressions are characteristic of written language (Biber et al., 1999), Thai learners displayed both written

and spoken linguistic features in their written discourse. However, the consistent use of a single tone is required for successful writing (Petch-Tyson, 1998).

(19) They can work in the department store **such as** BigC and Lotus, the restaurant **such as** KFC and McDonalds and the convenient store **such as** Seven-Eleven. (THA)
(20) **For example**, students spend their free time to work in their part-time job. (THA)

4.2.6 Indonesian

As Figure 2 shows, Indonesian learners underused all five interactional resources (viz., hedges, boosters, attitude markers, engagement markers, and self-mentions) and one interactive resource (viz., frame markers). It is more difficult for language teachers to identify underused patterns than to detect overused and misused patterns in their learners' writings. However, by utilizing the systematic information obtained from corpora, language teachers can explore the linguistic features not used by their learners due to proficiency, learner strategies, or cultural conventions. Besides, from the standpoint of contrastive rhetoric, it is interesting to qualitatively investigate the linguistic features used as a substitution for the underused features. For example, in Indonesian learners' essays, the frequent use of third person subjects, such as *part-time job(s)* and *college student(s)*, leads to the underuse of first person pronouns as self-mentions and second person pronouns as engagement markers.

(21) A **part-time job** can help you pay for living expenses, including food, books, gas, and clothes. (IDN)
(22) **College students** should have a part time job. (IDN)

One advantage of the corpus-based approach is the availability of a global description of learners' performances through the examination of a wide range of linguistic features. However, qualitative analysis is essential to unlock the full potential of corpus-based analysis (McEnery & Wilson, 2001). It can offer a rich and detailed perspective on the interlanguage, and complement quantitative analysis that can offer statistically reliable and generalizable results.

5. Conclusion

The purpose of the present study was to investigate differences of rhetorical

preferences in L2 writings, and identify discourse devices that can distinguish different L1 groups. The results indicate that there is a substantial difference in the use of metadiscourse markers between East Asian groups (viz., Chinese, Japanese, Korean, and Taiwanese) and Southeast Asian groups (viz., Indonesian and Thai). Furthermore, each learner group displayed the specific characteristics of metadiscourse, which offer suggestions for improving L2 learners' writings. These findings help teachers of academic writing correct their students' written compositions in English as well as assist language learners in becoming aware of rhetorical conventions in English academic writing.

The present study demonstrates that the methodological approach of combining comparable learner corpora, contrastive rhetorical analysis, and multivariate statistical methods is effective in exploring L2 metadiscourse. Corpus-based analysis enables linguists to capture stylistic deviations that are not typically identified in traditional discourse analysis since it can provide the frequency-based information on discourse devices in different corpora. The statistical information is crucial for understanding the pertinence of metadiscourse markers that have been outside the scope of traditional error analysis owing to the lack of clear rules for the usage. In particular, the knowledge about co-occurrence patterns is instrumental in mastering some types of metadiscourse markers that demonstrate its maximum rhetorical effect when used in combination with other markers in an appropriate manner. More detailed contrastive analysis of metadiscourse can reveal the relationships between learners' L1 and L2 performances. Additionally, the influence of teaching materials and teaching methods should be considered to unveil the causes of overuse and underuse found in learners' written productions.

Acknowledgments

Earlier versions of the present study were reported at the Language in Focus 2015 Conference and the Symposium of Second Language Writing 2015. This work was supported by Grants-in-Aid for Scientific Research Grant Numbers 26770205.

References

Ädel, A. (2006). *Metadiscourse in L1 and L2 English.* Amsterdam: John Benjamins.

Aijmer, K. (2002). Modality in advanced Swedish learners' written interlanguage. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 55-76). Amsterdam: John Benjamins.

Altenberg, B., & Tapper, M. (1998). The use of adverbial connectors in advanced Swedish learners' written English. In S. Granger (Ed.), *Learner English on computer* (pp. 80-93). London: Longman.

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers.* Amsterdam: John Benjamins.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use.* Cambridge: Cambridge University Press.

Biber, D., Johanson. S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English.* Harlow: Pearson Education.

Bickner, R., & Peyasantiwong, P. (1988). Cultural variation in reflective writing. In A. C. Purves (Ed.), *Writing across languages and cultures: Issues in contrastive rhetoric* (pp. 160-176). Newbury Park: Sage.

Bley-Vroman, R. (1983). The comparative fallacy in interlanguage studies: The case of systematicity. *Language Learning, 33,* 1-17.

Conner, U. (1996). *Contrastive rhetoric: Cross-cultural aspects of second-language writing.* Cambridge: Cambridge University Press.

Conrad, S. (2002). Corpus linguistic approaches for discourse analysis. *Annual Review of Applied Linguistics, 22,* 75-95.

Crismore, A., Markkanen, R., & Steffensen, M. (1993). Metadiscourse in persuasive writing: A study of texts written by American and Finnish university students. *Written Communication, 10(1),* 39-71.

Dahl, T. (2004). Textual metadiscourse in research articles: A marker of national culture or of academic discipline? *Journal of Pragmatics, 36,* 1807-1825.

Glynn, D. (2014). Correspondence analysis: An exploratory technique for identifying usage patterns. In D. Glynn & J. A. Robinson (Eds.), *Corpus methods in cognitive semantics: Quantitative studies in polysemy and synonymy* (pp. 443-485). Amsterdam: John Benjamins.

Granger, S. (1996). From CA to CIA and back: An integrated contrastive approach to computerized bilingual and learner corpora. In K. Aijmer, B. Altenberg, & M. Johansson (Eds.), *Languages in contrast: Papers from a Symposium on Text-based Cross-linguistic Studies, Lund 4-5 March 1994* (pp. 37-51). Lund: Lund University Press.

Granger, S., & Rayson, P. (1998). Automatic profiling of learner texts. In S. Granger (Ed.), *Learner English on computer* (pp. 119-131). London: Longman.

Hinkel, E. (2002). *Second language writers' text: Linguistic and rhetorical features.* Mahwah: Lawrence Erlbaum Associates.

Hong, H., & Cao, F. (2014). Interactional metadiscourse in young EFL learner writing: A corpus-based study. *International Journal of Corpus Linguistics, 19(2),* 201-224.

Hyland, K. (1998). Exploring corporate rhetoric: Metadiscourse in the CEO's letter. *The Journal of Business Communication, 35,* 224-246.

Hyland, K. (1999). Talking to students: Metadiscourse in introductory coursebooks. *English for Specific Purposes, 18,* 3-26.

Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing.* London: Longman.

Hyland, K. (2004). Disciplinary interactions: Metadiscourse in L2 postgraduate writing. *Journal of Second Language Writing, 13,* 133-151.

Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing.* New York: Continuum.

Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied Linguistics, 25,* 156-177.

Intaraprawat, P., & Steffensen, M. (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. *Learner Corpus Studies in Asia and the World, 1,* 91-118.

Kamimura, T., & Oi, K. (1998). Argumentative strategies in American and Japanese English. *World Englishes, 17(3),* 307-323.

Kim, L. C., & Lim, J. M. (2013). Metadiscourse in English and Chinese research article introductions. *Discourse Studies, 15(2),* 129-146.

Kobayashi, Y. (2011). Konparaburu kopasu wo motiita eisakubun to nihongosakubun no shiyogoi ni kansuru taisyoukenkyu [A contrastive study of the vocabulary use in English and Japanese writings with the use of comparable corpora]. In M. Nose (Ed.),

Taisyogengogaku wakate no kai sinpoziumu 2010 happyoronbunsyu [Proceedings of Young Researchers' Conference for Contrastive Linguistics 2010] (pp. 37-47). Nagoya: Sankeisha.

Kobayashi, Y. (in press). Heat map with hierarchical clustering: Multivariate visualization method for corpus-based language studies. *NINJAL Research Papers, 11*.

Lin, W. (2014). On metadiscoursal features of Chinese university students' oral English: A perspective from chunks. *US-China Foreign Language, 12(5),* 402-410.

McEnery, T., & Wilson, A. (2001). *Corpus linguistics: An introduction.* 2nd edition. Edinburgh: Edinburgh University Press.

Meyer, P. G. (1997). Hedging strategies in written academic discourse: Strengthening the argument by weakening the claim. In R. Markkanen & H. Shroder (Eds.), *Hedging and discourse: Approaches to the analysis of a pragmatic phenomenon in academic texts* (pp. 21-41). Berlin: Walter de Gruyter.

Oakes, M. P. (1998). *Statistics for corpus linguistics.* Edinburgh: Edinburgh University Press.

Petch-Tyson, S. (1998). Writer/reader visibility in EFL written discourse. In S. Granger (Ed.), *Learner English on computer* (pp. 107-118). London: Longman.

Reid, J. (1992). A computer text analysis of four cohesion devices in English discourse by native and nonnative writers. *Journal of Second Language Writing, 1(2),* 79-108.

Tan, H., & Eng, W. B. (2014). Metadiscourse use in the persuasive writing of Malaysian undergraduate students. *English Language Teaching, 7(7)*, 26-39.

Williams, J. (1981). *Style: Ten lessons in clarity and grace.* Boston: Scott Foreman.

Zamel, V. (1983). Teaching those missing links in writing. *ELT Journal, 37,* 22-29.