

A Contrastive Rhetoric Analysis of Metadiscourse Markers in Second Language Writing

Yuichiro Kobayashi

(Toyo University)

JWLLP 2015

- The earlier version of this presentation was reported at the Symposium on Second Language Writing 2015 in Auckland under the title of “Investigating metadiscourse markers in Asian Englishes: A corpus-based approach”

Introduction

- As **globalization** has increased intercultural and interlingual contacts, it is becoming important to understand the **diversity of World Englishes**
- Cultural differences in language has been the main topic of **contrastive rhetoric**, which identifies the writer's first language transfer to second language writing in terms of rhetorical strategy (Conner, 1996)
- Rhetorical preferences in first language can affect various aspects of language such as paragraph development (Bickner and Peyasantiwong, 1988), discourse development (Reid, 1992), and metadiscourse (Crismore, Markkanen, and Steffensen, 1993)

- Since the methodology of contrastive rhetoric is typically based on text linguistics, it is highly compatible with quantitative approach (Li, 2008)
- By using computerized learner corpora, linguists can obtain a large amount of frequency-based information on vocabulary, grammar, or discourse, which can be utilized for the comparisons among different learner groups

Related study

- Since the development of computerized learner corpora has flourished, **contrastive interlanguage analysis** has become a powerful framework in learner corpus research (Granger, 1996)
- Many previous studies on contrastive interlanguage analysis have utilized 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 corpus, the Louvain Corpus of Native English Essays (LOCNESS), which contains 324 thousand words of native writers' essays

- For example, 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 features characteristic of spoken language
- Aijmer (2002) also compared the frequencies of **modal devices** in **native speakers'** and **Swedish learners'** writings, and revealed the learners' overuse of all the modal categories examined in this paper

- The methodology of contrastive interlanguage analysis has been applied to the studies of Asian Englishes
- For instance, Abe, Kobayashi, and Narita (2013) compared native speakers and four learner groups (Hongkongese, Japanese, Korean, and Taiwanese), and identified linguistic features that can be used to discriminate between different learner groups and native speakers

Purpose

- The present study aimed to investigate **differences of rhetorical preferences in second language (L2)** writings among different first language (L1) groups
- This study compares the use of **metadiscourse markers** in L2 writings and identifies discourse devices that can be used to distinguish different L1 groups

Corpus data

- This 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 the present study is a subset from this corpus, including the written compositions of **six L1 groups** (Chinese, Indonesian, Japanese, Korean, Taiwanese, and Thai), which consists of expanding circle users of English in the viewpoint of World Englishes

- The subset analyzed in this study includes only writers with **B1 CEFR level**
- The writing conditions and learners' proficiency levels were strictly controlled for the comparison of these groups
- The subset contains essays written in response to **a single prompt**, namely “It is important for college students to have a part-time job” (Ishikawa, 2013, p. 95)

	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

Metadiscourse markers

- **Metadiscourse** is widely used term in current discourse analysis and contrastive analysis
- The present study is based on the framework of metadiscourse developed by **Ken Hyland**
- Hyland (2005) has 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)

- In this study, six learner groups were compared in terms of **the frequencies of nearly 500 types of metadiscourse markers** listed in Hyland (2005)

Category	Function	Examples
Interactive resources		
Transitions (TRA)	Express semantic relation between main clauses	<i>in addition, but, thus, and</i>
Frame markers (FRM)	Refer to discourse acts, sequences, or text stages	<i>finally, to conclude, my purpose here is to</i>
Endophoric markers (END)	Refer to information in other parts of the text	<i>noted above, see Fig, in section 2</i>
Evidentials (EVI)	Refer to source of information from other texts	<i>according to X, (Y, 1990), Z states</i>
Code glosses (COD)	Help readers grasp functions of ideational material	<i>namely, e.g., such as, in other words</i>
Interactional resources		
Hedges (HED)	Without writer's full commitment to proposition	<i>might, perhaps, possible, about</i>
Boosters (BOO)	Emphasize force or writer's certainty in proposition	<i>in fact, definitely, it is clear that</i>
Attitude markers (ATM)	Express writer's attitude to proposition	<i>unfortunately, I agree, surprisingly</i>
Engagement markers (ENG)	Explicitly refer to or build relationship with reader	<i>consider, note that, you can see that</i>
Self-mentions (SEM)	Explicit reference to author(s)	<i>I, we, my, our</i>

- The frequencies of metadiscourse markers were automatically calculated using **MDM tagger** (Kobayashi & Yamada, 2008) developed by myself
- Annotations for some metadiscourse markers were manually modified
 - { I }_SEM usually have rice in the morning { because }_TRA { my }_SEM family { think }_BOO it is good for health { and }_TRA { I }_SEM { prefer }_ATM rice to bread.
 - Recently, young people in Japan { tend to }_HED eat bread in the morning { or }_COD eat nothing.

Statistical methods

- **Heat map with hierarchical clustering** was used to investigate differences of metadiscourse between different L1 groups in the present study
- It is a powerful method for **visualizing multivariate data** such as large frequency tables for corpus-based linguistic analysis (Kobayashi, 2014)
- Its graphical representation offers **the statistical summary of complex co-occurrence patterns** of samples (e.g., learner groups) and variables (e.g., metadiscourse categories) as well as the original frequency information in the data

- Using the hierarchical clustering, the underlying meaningful patterns between samples and variables can be identified, and, in addition, the interpretation of those patterns can be validated with the heat map

Procedures

- This study calculated the frequencies of ten functional categories of metadiscourse markers in L2 writings of six learner groups
- Following the frequency counts, it quantitatively compared the frequencies using heat map with hierarchical clustering
- Finally, it qualitatively examined the usage examples of metadiscourse markers characteristic of each learner group

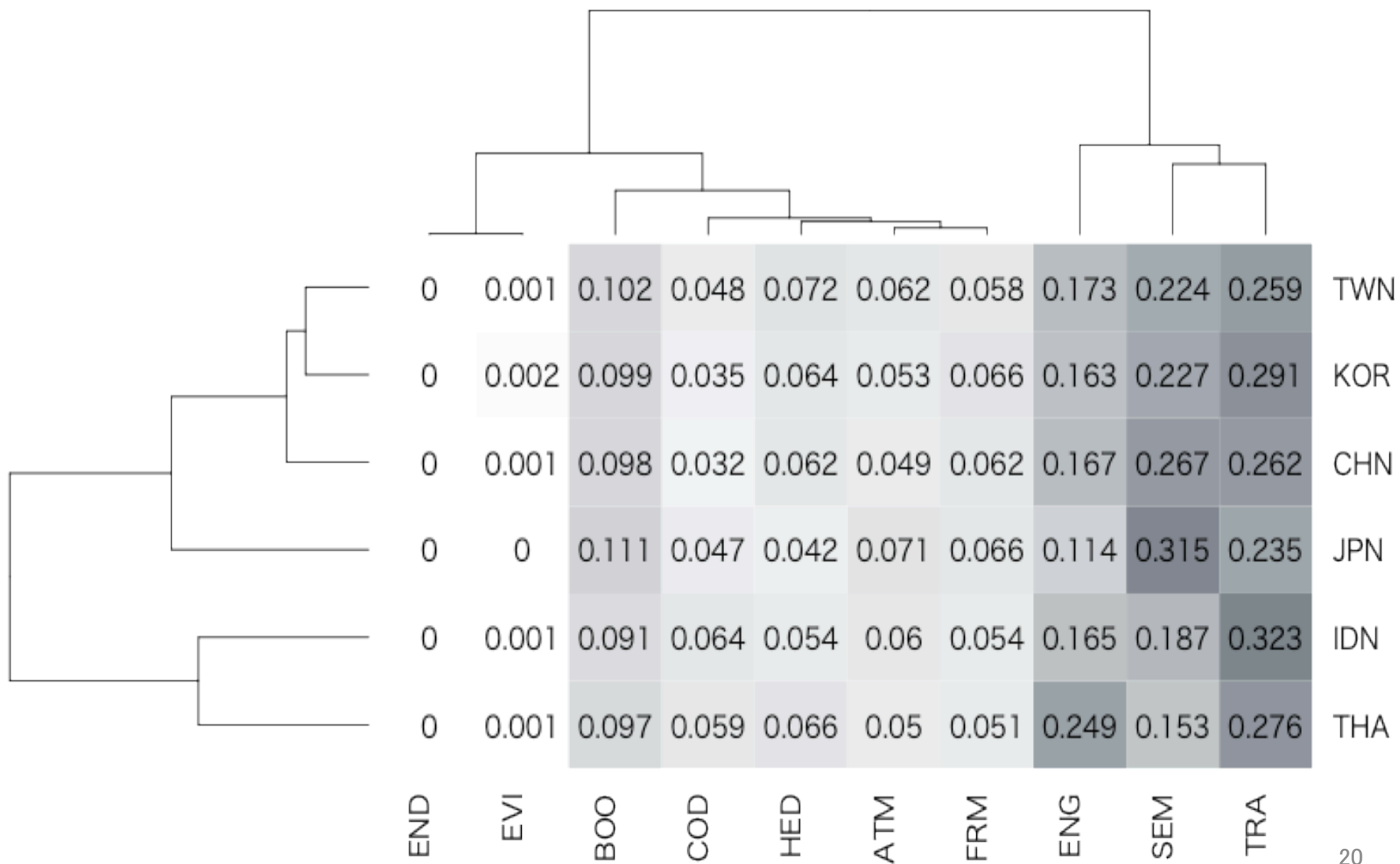
Results and discussions

- This study began by calculating the frequencies of metadiscourse categories in the writings of six learner groups
- The endophoric markers used in the writings are very low-frequent category, and all cell numbers in the row are zero because of the number of significant figures

	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
HED	0.062	0.054	0.042	0.064	0.072	0.066

- The next step was to investigate **the relationships between learner groups and metadiscourse categories** through heat map with hierarchical clustering
- The method displayed (a) the result of the clustering of learner groups, (b) the result of metadiscourse categories, and (c) the heat map generated from the permuted frequency table in two-dimensional space at the same time
- The complete linkage method on Euclidean distances (Divjak and Fieller, 2014) was used for the clusterings of learner groups and metadiscourse categories

- 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, metadiscourse categories in cells dark in color represent more frequent categories in that learner group, and categories in cells pale in color represent less frequent, as compared to in other groups
- Moreover, **relative frequencies were placed within each cell** in the heat map



Grouping of metadiscourse categories

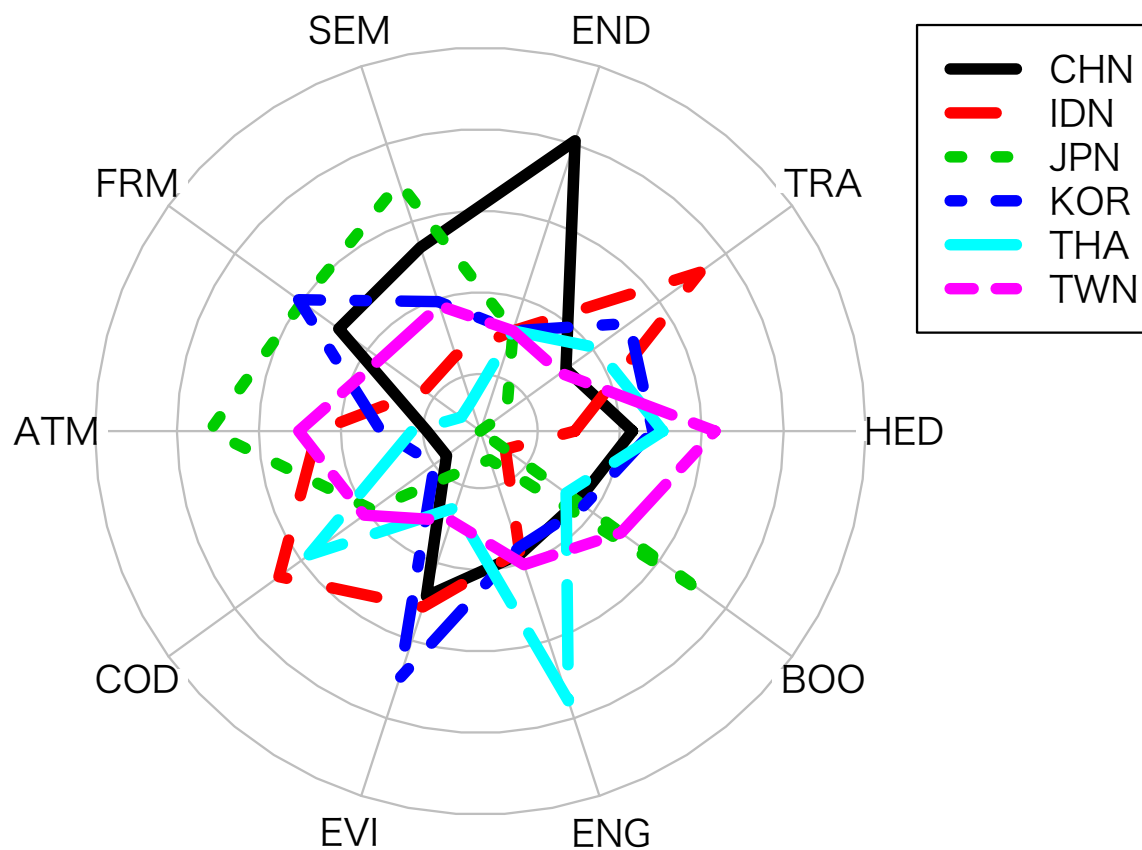
Grouping of learner groups

Frequency information

	END	EVI	BOO	COD	HED	ATM	FRM	ENG	SEM	TRA	
0	0.001	0.102	0.048	0.072	0.062	0.058	0.173	0.224	0.259	TWN	
0	0.002	0.099	0.035	0.064	0.053	0.066	0.163	0.227	0.291	KOR	
0	0.001	0.098	0.032	0.062	0.049	0.062	0.167	0.267	0.262	CHN	
0	0	0.111	0.047	0.042	0.071	0.066	0.114	0.315	0.235	JPN	
0	0.001	0.091	0.064	0.054	0.06	0.054	0.165	0.187	0.323	IDN	
0	0.001	0.097	0.059	0.066	0.05	0.051	0.249	0.153	0.276	THA	

- The clustering result of learner groups indicates that there is a substantial difference in the frequency patterns of metadiscourse markers between East Asian groups (Chinese, Japanese, Korean, and Taiwanese) and Southeast Asian groups (Indonesian and Thai)
- The result of metadiscourse categories showed high-frequency, middle-frequency, and low-frequency categories cluster together respectively

- Differences of metadiscourse among learner groups can be clearly shown in this radar chart



Japanese

- Japanese learners used more frequently self-mentions, boosters, frame markers, and attitude markers than other learner groups
- The most salient feature was **self-mentions** that refer to “the degree of explicit author presence in the text” (Hyland, 2005. p. 53)
- It is well-known that language learners are much more overtly present in their discourse than native speakers (Petch-Tyson, 1998), and Japanese learners are the most typical example

- They **write English as if they were speaking** since first person pronouns are linguistic features that characterize spoken language (Biber, Johanson, Leech, Conrad, and Finegan, 1999)
 - **I** am working at a convenience store near **my** home now. (JPN)
 - The experience was very important for **me**. (JPN)

- Another notable feature of Japanese learners was boosters that “head off conflicting views and express their certainty in what they say” (Hyland, 2005, p. 52)
- They overuse *think*, which follows *I*, and *of course* in the sentence-initial position
- The heavy use of these expressions is also due to influence of spoken language (Aijmer, 2002)
 - I think that college students should have a part time job. (JPN)
 - Of course we must study hard. (JPN)

Korean

- Korean learners made a significant use of **frame markers** that “signal text boundaries or elements of schematic text structure” (Hyland, 2005, p. 51)
- Overuse of these expressions was also reported by Tankó (2004) who examined the use of adverbial connectors in Hungarian university students’ essays
- It may result from “superficial attention” to logical forms (Intaraprawat and Steffensen, 1995, p. 271)
 - **First**, college students will get experiences about social job. (KOR)
 - **Second**, we can learn about style of living at working place. (KOR)

- Another notable feature of Korean learners was **evidentials** that are “metalinguistic representations of an idea from another source” (Thomas and Howes, 1994, p. 129)
- They refer to some surveys or newspapers to support their own claims
 - **According to** one survey, 73% of the students are planning to work at a part-time job in this summer vacation. (KOR)
 - **According to** a newspaper article college student’s work part-time wages are low. (KOR)

Chinese

- Chinese learners frequently used **endophoric markers** that “refer to other parts of the texts” (Hyland, 2005, p. 51)
 - For all these reasons mentioned **above**, it is important for college students to have a part-time job. (CHN)
 - I have the following reasons **below**. (CHN)

Taiwanese

- The prominent feature of Taiwanese learners was **hedges** that “indicate the writer’s decision to recognize alternative voices and viewpoints and so withhold complete commitment to a preposition” (Hyland, 2005, p. 52)
- This category is one of the most significant rhetorical devices in academic writing (Hyland, 2010)
- Good writers can use hedges for strengthening the argument by weakening the claim in their discourse (Meyer, 1997)

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

Thai

- The remarkable feature of Thai learners was **engagement markers** that “explicitly address readers, either to focus their attention or include them as discourse participants” (Hyland, 2005, p. 53)
- Thai learners used second person pronouns significantly more than other learner groups
 - Do you think this is a good idea? (THA)
 - You can help your parents to save their expenditure and you can save money for yourself. (THA)

Indonesian

- One of the characteristics of Indonesian learners was **transitions** that “help readers interpret pragmatic connections between steps in an argument” (Hyland, 2005, p. 50)
 - They can learn how to manage their time more appropriately, **because** they have to do their assignments and study for tests too. (IDN)
 - **Therefore**, part time job is important for college students. (IDN)

- Another characteristics was **code glosses** that “supply additional information, by rephrasing, explaining or elaborating what has been said, to ensure the reader is able to recover the writer’s intended meaning” (Hyland, 2005, p. 52)
- Indonesian learners displayed examples using **such as** or **for example**
 - They can do many jobs, **such as** a waiter, a computer mechanic, etc. (IDN)
 - **For example**, if they get part time job at restaurant, they get ability to service much people. (IDN)

Conclusion

- The purpose of this study was to investigate differences of rhetorical preferences in L2 writings, and to identify discourse devices that can be used to distinguish different L1 groups
- The findings suggest that there is **a substantial difference of the use of metadiscourse markers between East Asian groups** (Chinese, Japanese, Korean, and Taiwanese) **and Southeast Asian groups** (Indonesian and Thai)
- The prominent features of Japanese learners' metadiscourse were self-mentions, boosters, and attitude markers, and the notable features of Korean learners were frame markers and evidentials

- The salient feature of Chinese learners was endophoric markers, and the remarkable feature of Taiwanese was hedges
- Moreover, Thai learners frequently used engagement markers, and Indonesian learners made a significant use of transitions and code glosses
- More detailed analysis of metadiscourse can reveal the relationships between learners' L1 and L2 performances
- The present study contributes to the understanding of the nature and characteristics of variation in Asian Englishes

Further study

- To examine the influence of ...
 - learners' first language(s) and culture(s)
 - English textbooks used in junior and senior high schools
 - other factors in English teaching (e.g., classroom practice)

Summary of results

Learner groups	Metadiscourse characteristics
Japanese	Self-mentions (<i>I, my</i>), boosters (<i>think, of course</i>)
Korean	Frame markers (<i>first, second</i>), evidentials (<i>according to</i>)
Chinese	Endophoric markers (<i>above, below</i>)
Taiwanese	Hedges (<i>would, guess</i>)
Thai	Engagement markers (<i>you, your</i>)
Indonesian	Transitions (<i>because, therefore</i>), code glosses (<i>such as, for example</i>)

- **Contact:** Yuichiro Kobayashi (kobayashi0721@gmail.com)
- **Acknowledgments:** This work was supported by Grants-in-Aid for Scientific Research Grant Numbers 26770205 and 24320101