

PEDAGOGICAL STYLISTICS AND CONCEPT-BASED INSTRUCTION:  
AN INVESTIGATION INTO THE DEVELOPMENT OF VOICE IN THE ACADEMIC  
WRITING OF JAPANESE UNIVERSITY STUDENTS OF ENGLISH

by

Gary G. Fogal

A thesis submitted in conformity with the requirements  
for the degree of Doctor of Philosophy  
Graduate Department of Curriculum, Teaching and Learning  
Ontario Institute for Studies in Education, University of Toronto

© Copyright by Gary G. Fogal, 2015

UMI Number: 3715852

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 3715852

Published by ProQuest LLC (2015). Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against unauthorized copying under Title 17, United States Code



ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 - 1346

Pedagogical Stylistics and Concept-based Instruction:  
An Investigation into the Development of Voice in the Academic Writing of Japanese  
University Students of English

Gary G. Fogal  
Doctor of Philosophy

Department of Curriculum, Teaching and Learning  
University of Toronto  
2015

**Abstract**

This thesis study examines, through the lens of Vygotsky's sociocultural theory of mind, the influence and relationship of pedagogical stylistics (an analytic technique for examining text-based discourse features) and concept-based instruction (CBI) on academic writing performance and authorial voice construction in a high-stakes, essay writing context. Guided by CBI, 7 Japanese university-aged ESL students engaged in stylistics-based analyses of literary texts during a 3-week intensive course designed to prepare participants for the independent writing task of the TOEFL iBT. Participants were asked to consider how authors' lexical choices inform the semantic and pragmatic features of English literary texts associated with 5 target voice features: hedges, boosters, attitude markers, authorial self-mention, and direct reader references. Writing development was traced across 9 writing samples collected prior to, during, and after the intervention. The written compositions were analyzed and then compared, using non-parametric statistics, over time and with the writing of 3 students in a control group (that did not receive the intervention). Post-intervention stimulated recall and semi-structured interviews complemented the dataset. The analyses revealed: (a) statistically significant differences and large effect sizes across the written compositions, indicating improvements in learners' overall quality of writing ( $p = .002$ ;  $W = .5$ ) and in learners' overall quality of authorial voice ( $p = .002$ ;  $W = .51$ ); (b) authorial voice

development involved a nonlinear, iterative process of cognitive uptake derived from classroom-based affordances focused on stylistics; and (c) CBI-informed pedagogy mediated authorial voice development despite limitations in the teaching method. Based on these findings, instructors may consider employing stylistics by using CBI for improving students' overall writing and authorial voice in high-stakes writing contexts, implementing iterative learning environments for developing overall writing and authorial voice, focusing on writers' authorial voices in high-stakes writing contexts, and ensuring learners have sufficient opportunities to develop metalanguage in stylistics-informed contexts; researchers should examine the microgenetic development of authorial voice through the lens of a complex dynamic systems theory of SLA, expand the purview of L2 stylistics using empirically-based studies, and analyze how learners develop and use metalanguage in stylistic contexts.

## Acknowledgements

Completing my doctoral dissertation was made possible only with the guidance and support of many people, all of whom deserve my gratitude. Foremost I thank my supervisor, Alister Cumming. Alister is an ideal supervisor, and I am grateful and indebted to him for his mentorship. I hope to emulate his approach to supervision with my future students.

I also thank the other members of my thesis committee, Nina Spada and Enrica Piccardo. They provided me with constructive suggestions and their guidance helped to shape this research project for the better. I am also grateful to members of my examination committee, Christine Tardy, Sharon Lapkin, and Jeff Bale for their time and their fresh and insightful perspectives on my work. I also acknowledge the faculty and staff, as well as my colleagues in the Language and Literacies Education program at OISE: Thank you for providing an inspiring learning environment. And I would be remiss if I did not single out and thank Choongil Yoon who frequently provided sound advice. I also express my gratitude to the participants who volunteered their time to this study, and to the generous assistance of the Social Sciences and Humanities Research Council of Canada (grant number 752-2013-1291) for supporting this project.

This work also benefited greatly from the support of many friends and family members: Randy Brotto, Maria Diez, Bomin Hamlin, Dan Hamlin, Azusa McCourt, Kevin McCourt, and my parents and family in Windsor and Bangkok. Finally, to my lovely wife Lek and my darling children Airi and Risa: They bore the brunt of the hardships that inevitably come with such a large project and they provided the relief, patience, and encouragement I needed to sustain my efforts. Thank you.

## Table of Contents

Abstract.....	ii
Acknowledgements.....	iv
List of Tables.....	viii
List of Figures.....	x
List of Appendices.....	xii
Chapter 1. Introduction: Weaving Together Multiple Perspectives.....	1
1.1 Overview of the study and introduction of research questions.....	6
1.2 Key terms.....	8
1.2.1 Pedagogical stylistics.....	8
1.2.2 Voice.....	10
1.2.3 Concept-based instruction and systemic theoretical instruction.....	13
1.2.4 Microgenetic analysis.....	16
1.3 Motivation and interest for this thesis study.....	17
1.4 Overview of the thesis chapters.....	21
Chapter 2. Perspectives on Pedagogical Stylistics, Authorial Voice Development, and CBI.....	22
2.1 Pedagogical stylistics.....	22
2.1.1 Pedagogical stylistics for mediating L2 performance.....	23
2.1.2 Pedagogical stylistics for enhancing language awareness.....	25
2.1.3 Pedagogical stylistics for mediating and transferring learning and learning skills beyond language acquisition contexts.....	26
2.1.4 Underreporting and under-collecting of data in pedagogical stylistics research.....	29
2.2 Authorial voice development.....	33
2.2.1 Macrolevel research attending to voice development.....	34
2.2.2 Microlevel research attending to voice development.....	39
2.3 Concept-based instruction and systemic theoretical instruction.....	44
2.4 Other research domains.....	49
2.5 Summary.....	53
Chapter 3. Methods.....	55
3.1 Recruitment.....	56
3.2 Participant biodata.....	58
3.3 Instruments.....	59
3.4 Data collection and coding.....	63
3.4.1 Collecting and coding data from the TOEFL iBT independent task essay writing samples.....	64
3.4.2 Collecting and coding data regarding concept development.....	66

3.4.3 Collecting and coding self-reported verbal data .....	68
3.5 Data analyses .....	69
3.5.1 Research Question 1 (RQ1) .....	69
3.5.2 Research Question 2 (RQ2) .....	72
3.5.3 Research Question 3 (RQ3) .....	73
3.5.4 Research Question 4 (RQ4) .....	73
3.5.5 Research Question 5 (RQ5) .....	74
3.5.6 Self-reported verbal data .....	75
3.6 Summary .....	76
Chapter 4. Instructional Context .....	78
4.1 Researcher-cum-instructor .....	79
4.2 Biodata information .....	80
4.3 Instructional phase .....	83
4.4 Teaching with stylistics .....	90
4.5 Classroom activities .....	92
4.5.1 Week one (Days one through four) .....	92
4.5.2 Week two and three (Days five through twelve) .....	93
4.6 Feedback during the instructional phase .....	95
4.7 Narrative account of teaching-learning interactions .....	97
4.8 Summary .....	100
Chapter 5. Findings about Stylistics, Authorial Voice Development, and CBI .....	102
5.1 RQ1 .....	102
5.1.1 Summary of findings for RQ1 (part 1) .....	115
5.1.2 RQ1 (part two) .....	116
5.1.3 Summary of findings for RQ1 (part two) .....	118
5.2 RQ2 .....	119
5.2.1 Summary of findings for RQ2 .....	129
5.3 RQ3 .....	130
5.3.1 Summary of findings for RQ3 .....	134
5.4 RQ4 .....	135
5.4.1 Summary of findings for RQ4 .....	142
5.5 RQ5 .....	142
5.5.1 Summary of findings for RQ5 .....	144
5.6 Findings related to microgenetic processes .....	144
Chapter 6. Discussion .....	147
6.1 Summary of findings for RQ1 .....	147
6.2 Summary of findings for RQ2 .....	151
6.3 Summary of findings for RQ3 .....	152
6.4 Summary of findings for RQ4 .....	154

6.5 Summary of findings for RQ5 .....	156
6.6 Theoretical and pedagogical implications .....	156
6.6.1 Implications for authorial voice development .....	157
6.6.2 Implications for concept-based instruction .....	163
6.6.3 Implications for pedagogical stylistics .....	168
6.7 Recommendations for future research .....	173
6.8 Summary remarks .....	179
References .....	182
Appendices .....	214



## List of Tables

Table 1.1 Example of Concept-based Learning with Stylistics, Using the Literary Device Rhetorical Question.....	15
Table 3.1 Summary of Research Questions, Research Instruments, and Data Analyses .....	77
Table 4.1 Participant Biodata .....	82
Table 4.2 Instructional Procedures, Aims, and Rationale .....	84
Table 5.1 Summary of Means and Standard Deviations for Writing Scores and Voice Scores Over Time for Participants in the Intervention ( $n = 7$ ) .....	102
Table 5.2 Follow-up Pair-wise Tests Regarding Difference Over Time in the Overall Quality of Writing.....	104
Table 5.3 Follow-up Pair-wise Tests Regarding Difference Over Time in the Overall Quality of Voice in Writing.....	105
Table 5.4 Tally of Responses Regarding Mediating Tools and Rhetorical Feature Use ( $n = 7$ ) .....	106
Table 5.5 Summary of Comments Regarding the Role of Stylistics as a Mediating Tool for the Concept Development of the Target Rhetorical Features ( $n = 7$ ) .....	110
Table 5.6 Summary of Means, Standard Deviations, and Wilcoxon Signed-rank Test for Writing Scores and Voice Scores Sustained Over Time ( $n = 7$ ).....	117
Table 5.7 Summary of Means and Standard Deviations for Frequency of Use of Rhetorical Features Over Time ( $n = 7$ ) Across 49 Essays.....	119
Table 5.8 Correlations, Means, and Standard Deviations of the Overall Quality of Writing and the Frequency of Use of Each Rhetorical Feature of Voice Across 49 Essays .....	120
Table 5.9 Correlations, Means, and Standard Deviations of the Overall Quality of Voice in Writing and the Frequency of Use of Each Rhetorical Features of Voice Across 49 Essays .....	121
Table 5.10 Summary of Responses Regarding Participants' Interpretations of the Overall Quality of Voice in their Posttest Writing Task ( $n = 7$ ) .....	132
Table 5.11 Correlations, Means, and Standard Deviations of the Frequency of Use of Each Target Rhetorical Feature of Voice and the Concept Development (CD) of Each Target Rhetorical Feature of Voice Across 35 Writing Tasks .....	136
Table 5.12 Correlations, Means, and Standard Deviations of the Frequency of Use of Each Target Rhetorical Feature of Voice and the Concept Development (CD) of Each Target Rhetorical Feature of Voice Across 5 Speaking Tasks .....	137
Table 5.13 Means and Standard Deviations for Overall Quality of Writing Scores and Overall Quality of Voice in Writing Scores for the Control Group ( $n = 3$ ) and the Intervention Group ( $n = 7$ ) Across 20 Essays.....	143

Table 5.14 Mann-Whitney Analyses for Differences in the Quality of Writing and the Quality of Voice in Writing Between the Control Group (n = 3) and the Intervention Group (n = 7) Across the First Writing Sample and the Delayed Posttest, Across 20 Essays .....	143
Table 5.15 Data Documenting Microgenetic Changes in Tom's Development.....	145

## List of Figures

<i>Figure 2.1.</i> Cyclical pattern of cognitive learning processes for authorial voice development.....	42
<i>Figure 4.1.</i> Regulatory scale–implicit (strategic) to explicit .....	97
<i>Figure 5.1.</i> Radar graph depicting Maru’s uses of each target rhetorical feature of voice. .	122
<i>Figure 5.2.</i> Radar graph depicting Anne’s uses of each target rhetorical feature of voice. .	123
<i>Figure 5.3.</i> Radar graph depicting Tom’s uses of each target rhetorical feature of voice. ..	124
<i>Figure 5.4.</i> Radar graph depicting Anny’s uses of each target rhetorical feature of voice. .	125
<i>Figure 5.5.</i> Radar graph depicting Pi’s uses of each target rhetorical feature of voice.....	126
<i>Figure 5.6.</i> Radar graph depicting Am’s uses of each target rhetorical feature of voice. ....	127
<i>Figure 5.7.</i> Radar graph depicting Angela’s uses of each target rhetorical feature of voice.....	128
<i>Figure 5.8.</i> Group means depicting the use of each target rhetorical feature of voice across writing tasks ( $n = 7$ ).....	129
<i>Figure 5.9.</i> Relationship between mean concept development (CD) scores and frequency of use of hedges ( $n = 7$ ). .....	137
<i>Figure 5.10.</i> Relationship between mean concept development (CD) scores and frequency of use of boosters ( $n = 7$ ).....	138
<i>Figure 5.11.</i> Relationship between mean concept development (CD) scores and frequency of use of attitude markers ( $n = 7$ ) .....	138
<i>Figure 5.12.</i> Relationship between mean concept development (CD) scores and frequency of use of authorial self-mention ( $n = 7$ ).....	139
<i>Figure 5.13.</i> Relationship between mean concept development (CD) scores and frequency of use of direct reader references ( $n = 7$ ).....	139
<i>Figure N1.</i> Change over time regarding overall quality of writing ( $n = 7$ . .....	237
<i>Figure N2.</i> Change over time regarding overall quality of voice ( $n = 7$ ).....	237
<i>Figure O1.</i> Effects of the intervention sustained over time based on the posttest and the delayed posttest.....	239
<i>Figure P1.</i> Relationship between the frequency of use of hedges and essay scores .....	240
<i>Figure P2.</i> Relationship between the frequency of use of boosters and essay scores.....	240
<i>Figure P3.</i> Relationship between the frequency of use of attitude markers and essay scores .....	241
<i>Figure P4</i> Relationship between the frequency of use of authorial self-mention and essay scores.....	241

<i>Figure P5</i> Relationship between the frequency of use of direct reader references and essay scores .....	242
<i>Figure Q1.</i> Relationship between the frequency of use of hedges and voice scores .....	243
<i>Figure Q2.</i> Relationship between the frequency of use of boosters and voice scores .....	243
<i>Figure Q3.</i> Relationship between the frequency of use of attitude markers and voice scores .....	244
<i>Figure Q4.</i> Relationship between the frequency of use of authorial self-mention and voice scores .....	244
<i>Figure Q5.</i> Relationship between the frequency of use of direct reader references and voice scores .....	245
<i>Figure R1.</i> Relationship between writing scores and voice scores across writing tasks and across participants ( $n = 7$ ).....	246
<i>Figure S1.</i> Relationship between the frequency of use of hedges in writing tasks and the concept development (CD) of hedges in writing tasks across participants .....	247
<i>Figure S2.</i> Relationship between the frequency of use of boosters in writing tasks and the concept development (CD) of boosters in writing tasks across participants .....	247
<i>Figure S3.</i> Relationship between the frequency of use of attitude markers in writing tasks and the concept development (CD) of attitude markers in writing tasks across participants.....	248
<i>Figure S4.</i> Relationship between the frequency of use of authorial self-mention in writing tasks and the concept development (CD) of authorial self-mention in writing tasks across participants .....	248
<i>Figure S5.</i> Relationship between the frequency of use of direct reader references in writing tasks and the concept development (CD) of direct reader references in writing tasks across participants .....	249
<i>Figure T1.</i> Relationship between the frequency of use of hedges in writing tasks and the concept development (CD) of hedges in speaking tasks across participants.....	250
<i>Figure T2.</i> Relationship between the frequency of use of boosters in writing tasks and the concept development (CD) of boosters in speaking tasks across participants .....	250
<i>Figure T3.</i> Relationship between the frequency of use of attitude markers in writing tasks and the concept development (CD) of attitude markers in speaking tasks across participants.....	251
<i>Figure T4.</i> Relationship between the frequency of use of authorial self-mention in writing tasks and the concept development (CD) of authorial self-mention in speaking tasks across participants.....	251
<i>Figure T5.</i> Relationship between the frequency of use of direct reader references in writing tasks and the concept development (CD) of direct reader references in speaking tasks across participants.....	252

## List of Appendices

Appendix A. Recruitment Poster .....	215
Appendix B. Information Letter and Consent Form for Outside Raters .....	216
Appendix C. Biodata Questionnaire .....	219
Appendix D. TOEFL iBT independent writing task scoring rubric .....	222
Appendix E. Voice Scoring Rubric .....	223
Appendix F. Tally Framework for Measuring Frequency of Use of Rhetorical Features ...	228
Appendix G. Stylistic Analysis Worksheet .....	229
Appendix H. Script for Combined Stimulated Recall & Semi-structured Interview Protocol .....	231
Appendix I. TOEFL iBT Independent Writing Task Essay Prompts .....	233
Appendix J. Coding Framework for Concept Development of Rhetorical Features of Voice in TOEFL iBT Independent Writing Task and Stylistic Worksheets .....	234
Appendix K. Coding Framework for Concept Development of Rhetorical Features of Voice in Audio Recordings of Classroom Discussions .....	235
Appendix L. Orienting Basis of Action .....	236
Appendix M. Visual Aid II .....	237
Appendix N. Line Graphs for RQ1 .....	238
Appendix O. Line Graph for RQ1 (part two) .....	239
Appendix P. Scatterplots for RQ2: Writing Scores and Rhetorical Feature Use .....	240
Appendix Q. Scatterplots for RQ2: Voice Scores and Rhetorical Feature Use .....	243
Appendix R. Scatterplot for RQ3: Writing Scores and Voice Scores .....	246
Appendix S. Scatterplots for RQ4: Rhetorical Feature Concept Development Scores and Writing Tasks .....	247
Appendix T. Scatterplots for RQ4: Rhetorical Feature Concept Development Scores and Speaking Tasks .....	250
Appendix U. Transcription Conventions .....	253
Appendix V. Summary of Research Findings Organized by Research Question .....	254

## Chapter 1

### Introduction: Weaving Together Multiple Perspectives

This thesis study examines the influence and relationship of pedagogical stylistics (an analytic technique for examining text-based discourse features) and concept-based instruction (CBI) on academic writing performance and authorial voice construction in a high-stakes, essay writing context. Viewed through the lens of Vygotsky's sociocultural theory of mind, the research examined three distinct areas of second language (L2) research and practice: pedagogical stylistics, authorial voice development, and CBI. In this introduction I sketch the background of each dimension and briefly introduce gaps in the literature that would benefit from further inquiry. In doing so I highlight the rationale for including these areas of research in this study, and then I demonstrate how these three domains fuse into a cohesive research project.

The potential for pedagogical stylistics (defined in Subsection 1.2.1) to advance L2 learning—and by extension L2 academic writing—is underdeveloped (Bellard-Thomson, 2010, 2011; Carter, 2007, 2010; Edmondson, 1997; Fialho, Miall, & Zyngier, 2012; Fogal, 2015; Hall, 2007, 2014; Paran, 2008; Verdonk, 2002; Zyngier & Fialho, 2010). This lack of understanding has been triggered by the historical trajectory and focus of stylistics-based research (I use stylistics and pedagogical stylistics interchangeably in this thesis). With rare exceptions, the primary aim of research in this domain has focused on how best to raise students' awareness of language use primarily (though not exclusively—see Carter & Nash, 1990) in literary texts rather than focusing on stylistics as a tool for mediating students' L2 proficiency (Clark & Zyngier, 2003). That stylistic analyses aimed at raising language

awareness—one of the primary aims of a stylistics-informed pedagogy (Clark & Zyngier, 2003)—would lead to students’ improved L2 performance was “often assumed” (Clark & Zyngier, 2003, p. 341), but rarely demonstrated systematically, as Bellard-Thomson, Carter, Hall, and many others have noted. The resultant lack of attention, then, has compounded, and research to date into stylistics has focused almost exclusively on how best to employ stylistics as a teaching tool with little attention to, in the L2 context, the effects of this approach on L2 proficiency gains (Fogal, 2015; Hall, 2007, 2014; Paran, 2008). While this shortcoming has been documented extensively in the literature, Paran’s account best underscores the gravity of this problem for the future of stylistics in L2 contexts: “If stylistics wishes to capture an important place in [L2] language learning it will have to address the issues which language teachers and learners are preoccupied with, and will have to demonstrate the usefulness of the approach to language teaching” (p.487) that has as its end goal, mediating L2 development.

To date, numerous studies (Al-Jarf, 2007; Allen, 2009; Badran, 2012; Davies, 1998; Fogal, 2010; Lin, 2010; Paesani, 2006; Plummer & Busse, 2006; Saugera, 2011; Timucin, 2001; Warner, 2012; Yañez Prieto, 2010; Zerkowitz, 2012) have addressed Paran’s (2008) call to examine the efficacy of stylistics for mediating L2 development, and done so with some degree of success. However, many of these studies—as I highlighted in a review synthesis prepared in tandem with this thesis study (Fogal, 2015)—lack empirical data and corresponding firm evidence to validate stylistics’ effectiveness in L2 learning contexts and, as I later demonstrate (see Section 2.1), in L2 academic writing classrooms. At present, studies are needed that can validate the effectiveness of stylistics for mediating students’ L2 development.

Numerous studies have investigated authorial voice development in L2 writing contexts but as a concern separate from pedagogical stylistics (in Subsection 1.2.2 I define “voice” and I outline how voice is operationalized in this study). For example, research to date has expanded understandings of voice development in university writing contexts (Ramanathan & Kaplan, 1996; Zhao, 2010) and high-stakes writing contexts (Jeffrey, 2009; Zhao, 2010, 2013). The ubiquitous presence of voice as a criterion in instructional and assessment materials in both of these writing contexts suggests that most English L2 learners attempting to gain access to or succeed in English-medium universities will be required to produce representations of voice in their writing, in one form or another. Although recent studies have successfully widened understandings of voice development in these contexts, there remain avenues for further studying voice in English L2 academic writing. Accordingly, I argue that this particular construct of academic writing merits continued attention.

Some recent studies (e.g., Bayyurt, 2010, Hirvela & Belcher, 2001, Kim, Baba, & Cumming, 2006, and Tardy, 2009) have provided a clear basis for understanding how learners negotiated macrolevel, socio-cultural factors such as personal histories and (typically newly-entered) discourse communities in the process of constructing authorial voice. For example, Kim et al. (2006) documented the macrolevel struggles of L2 learners acculturating to new discourse spaces by examining learners’ impressions of voice construction and their perceived expectations of their discourse communities. These and other findings (reviewed in Section 2.2) underscore the importance of macrolevel influences on voice construction and provide insights into how learners have acculturated to and developed their authorial voices. However, despite the wealth of knowledge accrued from



these studies, understanding voice construction through the lens of Vygotsky's sociocultural theory of mind remains incomplete.

Viewed through Vygotsky's theory, understanding how learners negotiate the complex task of constructing an authorial voice requires two avenues of study: macrolevel studies of learners and their socio-cultural learning contexts; and microlevel, microgenetic studies (defined in Subsection 1.2.4) that reveal the cognitive processes informing voice development. As I have already noted, numerous studies have been successful at outlining macrolevel concerns. However, microlevel understandings of voice construction are still needed to complement present macrolevel findings—with the end goal of more accurately interpreting voice development. To date, limited research has examined this microgenetic context. For example, P. Chang (2010) documented the types of thinking that inform voice development, and in a study I prepared based on the pilot phase of this thesis study (Fogal, submitted) I documented the sequence of cognitive processing as learners engaged target language features that inform voice construction. While both of these studies take important and necessary first steps toward understanding voice development through a microgenetic lens, more in-depth, systematic studies are needed to reveal how learners engage this complex construct. Within the context of this thesis study, authorial voice provides a relevant—and workable—construct for examining the effectiveness of pedagogical stylistics at informing L2 development.

Thus far I have highlighted that authorial voice provides a writing construct for better understanding stylistics as a mediating tool for L2 writing development. The remaining dimension that I now introduce for this thesis study is CBI.

The last decade has witnessed a surge in classroom research on CBI (Brooks, Swain, Lapkin, & Knouzi, 2010; Ferreira, 2005; Ferreira & Lantolf, 2008; Gánem-Gutiérrez & Harun, 2011; Garcia, 2012; Johnson, 2008; Kim, 2013; Knouzi, Swain, Lapkin, & Brooks 2010; Lai, 2012; Lapkin, Swain, & Knouzi, 2008; Lee, 2012; Negueruela, 2003; Negueruela & Lantolf, 2005; Swain, Lapkin, Knouzi, Suzuki, & Brooks, 2009; Twyman, Ketterlin-Geller, McCoy, & Tindal, 2003; van Compernelle, 2011, 2012; Yañez-Prieto, 2008, 2010). Concept-based instruction is a pedagogical manifestation of Vygotsky's theoretical work on spontaneous and scientific concepts in human development (Vygotsky, 1986). Concept-based instruction (defined in detail in Subsection 1.2.3) is an attempt to mediate learner development by introducing "scientific" concepts (in this thesis study, operationalized as five rhetorical voice features, outlined in Subsection 1.2.2) that build on pre-existing, "spontaneous" concepts that learners already possess (Swain, Kinnear, & Steinman, 2011; Wertsch, 1991). The studies listed above have demonstrated the efficacy of a CBI approach to L2 teaching. For example, Swain and colleagues have repeatedly documented the effectiveness of languaging (an integral feature of CBI-informed pedagogies) for mediating L2 development, while Negueruela and Lantolf have demonstrated the advantages of a CBI approach to developing grammatical competency. While these studies are extensive and provide a solid foundation for understanding CBI-informed pedagogies and their influence on mediating L2 development, only limited studies have examined CBI in L2 academic writing contexts (Ferreira, 2005; Ferreira & Lantolf, 2008; Johnson, 2008; Yañez-Prieto, 2008, 2010). For example, Johnson documented that learners internalized genre-based scientific concepts via CBI-informed mediation that enabled their L2 academic writing improvement. In two other studies, Ferreira, and Ferreira and Lantolf both found that

learners' genre writing had improved (although they were unable to discern the cause of the change), that learners' had problems orienting to a CBI-learning style, and that learners developed a conceptual understanding of genre but were unable to consistently apply this understanding to improve their writing. These studies are particularly insightful regarding how learners negotiate genre-based writing, the problems learners incur in a CBI-informed pedagogy, and how CBI can be employed to aid learners in similar genre-based studies. However, while these studies are informative and relevant, this research domain would benefit from continued research on other CBI and L2 writing contexts. To this end, this thesis study expands the understanding of CBI's usefulness to include concept development and voice construction in one L2 writing context.

To summarize, for the purpose of this thesis study I have woven together pedagogical stylistics, voice development, and CBI in the following manner. During the intervention phase of this thesis study CBI acted as a pedagogical guide for applying stylistics-informed analytic techniques to literary text analyses, with the aim of developing learners' overall quality of voice in their academic writing. This framework served as a backdrop for analyzing the effectiveness of stylistics as a teaching tool, the microgenetic processes informing voice development, and the usefulness of CBI for concept development in one, L2 writing context.

### **1.1 Overview of the study and introduction of research questions**

The present study sought to understand better the role of pedagogical stylistics as a tool for mediating authorial voice development in English L2 learners and, in tandem, to document a descriptive, microgenetic report of the cognitive processes implicated during authorial voice construction. To achieve these goals, I addressed the general question: *How,*

*and to what degree (if any), do qualities of voice improve in Japanese university students' writing in English when pedagogical stylistics, implemented through CBI, is employed as a mediational tool in the academic writing context?* I posed five research questions to guide this inquiry:

- 1) Is there an increase, over the intervention period, in participants' (a) overall quality of writing, and (b) overall quality of voice in writing? Is any increase in the quality of writing and the quality of voice sustained after the intervention?
- 2) What is the relationship between the frequency of use of each target rhetorical feature of voice, and (a) the overall quality of writing and (b) the overall quality of voice?
- 3) What is the relationship between (a) the overall quality of writing and (b) the overall quality of voice in writing?
- 4) What is the relationship between (a) the frequency of use of each target rhetorical feature of voice and (b) the concept development of each rhetorical feature of voice?
- 5) Are there differences in (a) the quality of writing and (b) the quality of voice in writing between the control group and the course participants?

As described in Chapter 3, I employed a sequential mixed-method research design to help answer these research questions. Ten Japanese, English as a second language (ESL) learners were recruited from the local community to participate in this study. Seven of these students engaged in a three-week intervention period where I instructed them on how to employ stylistic analyses of literary texts to improve their authorial voice while producing written compositions in response to the independent writing task of the Test of English as a Foreign Language internet-based Test (TOEFL iBT, Educational Testing Service, 2008). The remaining three participants formed a control group that only performed two tasks: The

control completed a writing task that established their baseline writing scores, and then they completed the posttest writing task. Both tasks were performed alongside the participants in the intervention group, using the same writing prompts as the intervention group. Data collected from the intervention period was paired with post-intervention, stimulated recall and semi-structured interviews done individually with each intervention-group participant. I analyzed the dataset to provide insights into the effectiveness of stylistics as teaching tool (and CBI as a pedagogical model) in this context and examined microgenetic, cognitive features that informed voice construction. This thesis study documents these findings, highlighting insights relevant to (and between) theory and practice regarding pedagogical stylistics, authorial voice development, and CBI.

## 1.2 Key terms

Below I define key concepts employed in this study.

### 1.2.1 Pedagogical stylistics

Following McIntyre (2011) and Short (1996), I define pedagogical stylistics as the application of a systematic, analytic technique for exploring language use and language meaning in literary and non-literary texts in (for the purpose of this study) L2 classroom-based contexts. This definition of pedagogical stylistics emphasizes the unique elements of literary analysis that stylistics has established through a long history of literature-use in ESL and English as a Foreign Language (EFL) contexts (see Belcher & Hirvela, 2000, Hirvela, 2004, Khatib, Rezaei & Derakhshan, 2011, and Paran, 2008, for extensive discussions). Stylistics may be best understood in contrast to reader-response approaches to literary studies. Such approaches to interpreting literary texts have long informed—and continue to inform—L2 classroom pedagogy (Belcher & Hirvela, 2000; Harding, 2014; Hirvela, 2004;

Khatib et al., 2011; Paran, 2008; Rosenblatt, 1978) with a specific focus on L2 composition (Belcher & Hirvela, 2000; Crain, 1993; Gajdusek & van Dommelen, 1993; Hirvela, 2001, 2004; Vandrick, 2003) wherein learners are (typically) engaged in personal response tasks (Hirvela, 2004; Vandrick, 2003). These response tasks are predicated by macrolevel discourse structures such as theme, character, symbolism, or plot (Gajduske & van Dommelen, 1993; Vandrick, 2003), for example, and serve learners as they provide textual support for their initial reactions to a text (Harding, 2014). In contrast, stylistics substitutes these macrolevel discourse structures for analyses of microlevel structures such as lexis, semantics, or syntax (Bellard-Thomson, 2011; Verdonk, 2013). Moreover, the process of interpretation is also inverted. Stylistic analyses begin by assuming that language precedes content (Carter, 1997), thus foregrounding language use and language meaning rather than a reader's response to a given text: "The basic assumption is that literature is made from and with language, that language is the medium of literature and that beginning with the very textuality of the text is a secure foundation for its interpretation" (Carter, 2010, p. 58). To achieve this focus on textuality, stylistic analyses begin with a systematic analysis of a text that precedes a final interpretation (Short, 1996). A differentiating factor of stylistics, then, is that stylistics requires that language learners analyze texts systematically to uncover how language use generates meaning. Only then are data collected from such analyses applied by readers to aid in interpreting a text. The systematic approach informing stylistics goes beyond what Hirvela (2004) described as direct models of reading for writing (i.e., mining texts for language resources and writerly reading) and as such helps differentiate stylistics from other approaches to literary-text analysis (of which Khatib et al., 2011 provide an extensive list) that foreground the reader rather than the language.

The practice of applying stylistics, although grounded in seminal works such as Short (1996) and categorized in Bellard-Thomson (2011), does not follow a single pedagogical model. Purposively, then, stylisticians have concluded, “it is difficult if not impossible to define the content of stylistics courses with any certainty” (Bellard-Thomson, 2011, p. 55). However, while stylistic techniques may vary, they are typically systematic (Short, 1996). Exploratory techniques employed during stylistic analyses focus on (but are not limited to) examining graphology, phonology, lexis, syntax, semantics, and pragmatics (Bellard-Thomson, 2011; Verdonk, 2013) with attention to either one or a multitude of these dimensions. Widely employed formats for analysis include check sheets (Short, 1996) that focus readers’ attention on target language features embedded in, for example, pragmatics or syntactic structures. The aim of this technique is to guide learners systematically to collect and employ linguistic evidence to support their interpretations of a text with the shared aim of assisting language learners in becoming “linguistically aware readers who can perceive the qualities of language which are manipulated for particular effects” (Clark & Zyngier, 2003, p. 342). Accordingly, stylisticians argue that this perspective translates into increased textual understanding (Bellard-Thomson, 2011) and in some cases, as this thesis study has attempted to establish, increased language proficiency.

### 1.2.2 Voice

In their review article, “Voice in the Context of Literacy Studies”, Sperling and Appleman (2011) observed the following: “Voice is as complicated as the self it is assumed to evoke and is as socially situated and culturally embedded as the self is understood to be” (p. 82). As this comment suggests, the construct of voice in writing is a slippery and complex

term. For example, in first language (L1) contexts voice has been characterized as an expression of the self, often realized through expressive writing as a tool for self-discovery (Elbow, 1968, 1994, 1999; Graves, 1994). Increasingly however, and in response to culturally and linguistically diverse learning contexts, attempts at defining voice through the lens of Western individualism have been dismissed (Bowden, 1999; Hirvela & Belcher, 2001; Matsuda, 2001; Ramanathan & Atkinson, 1999). Present understandings of voice are less monologic, evoking instead a Bhaktinian perspective (Hyland, 2008; Sperling & Appleman, 2011) that foregrounds history and culture as fundamental to the evolving concept of voice while still acknowledging that writers imbue voice with their personal linguistic choices (Matsuda, 2001; Zhao, 2013). Present day conceptualizations of voice in the L2 academic writing context, then—and the understanding of voice that informs this thesis study—are underpinned by the coexistence of the personal with the communal (Hewings, 2012; Hyland & Guinda, 2012; Matsuda, 2001; Prior, 2001) wherein constructing a dialogic voice is employed to engage varying discourse communities (Hyland, 2008; Matsuda, 2001).

Attempts to understand how L2 writers engage various discourse communities have spawned studies that address the role of authorial voice in negotiating this process (Hood, 2012; Hyland & Guinda, 2012; Ivanič & Camps, 2001, Matsuda & Jeffery, 2012; Matstuda & Tardy, 2007; Thomson, 2012; Tse, 2012). To this end, Hyland (2005, 2008) developed an interactionist model of voice that examines the author-reader interface. Hyland's model—which draws from his corpus-based research (Hyland, 1994, 1996, 1996a, 1999, 2004a) and from similar works (Bondi, 1999; Kuo, 1999; Thompson & Ye, 1991), and from which this thesis study draws heavily—is composed of two primary linguistic categories: stance



(comprised of hedges, boosters, attitude markers, and authorial self-mention) and engagement (comprised of direct reader references, directives, questions, knowledge reference, and asides) that highlight the linguistic features writers employ to construct voice and to engage a target discourse community. Zhao (2013) adapted this model to address the writing context of her study, suggesting that Hyland’s model provided a systematic framework for exploring how voice is constructed and manifest “through the use of both linguistic- and discourse-level elements” (Zhao, 2013, p. 203). Zhao refined this model to include only those voice features most likely to concern writers composing the argumentative writing task of the TOEFL. Zhao argued that the constraints of the TOEFL typically preclude the use of questions, knowledge references, and asides in test-taker compositions (see Coffin & Hewings, 2004 for an analysis of how similar constraints affect International English Language Testing System [IELTS] test-takers). Given that my study also examined responses to the independent writing task of the TOEFL test to examine voice, this thesis study draws from Zhao’s framework, and operationalizes the construct of voice as the use of hedges, boosters, attitude markers, authorial self-mention, and direct reader references. More specifically, hedges are linguistic tools that enable writers to distance themselves from a claim (e.g., perhaps, maybe, or tend to), while boosters enable writers to express certainty about a position (e.g., definitely, indeed, absolutely, or must). Attitude markers convey the writer’s attitude toward a claim, expressing “surprise, agreement, importance, [and] frustration” (Zhao, 2010, p. 43), for example. The remaining two discourse features examined in this study are linguistic tools employed to directly engage the reader. These tools include direct reader references (e.g., we, us, you, or our) and authorial self-mentions (e.g., I, me, or my) (see Hyland 2005, 2008 and Zhao, 2010, 2013 for detailed discussions

regarding these discourse features). While addressing voice in this manner reveals tensions between prescriptive accounts of voice atypical of recent L2 writing literature and more socio-culturally rooted attempts to understand authorial voice construction, the situated context of the intervention employed in this study and the learning goals of the participants merit this more narrowly construed operationalization of voice.

To summarize, in this thesis study I conceptualize voice in academic writing as an expression of an amendable, disciplinary self, realized through the linguistic interplay of the individual and the social wherein writers engage with discourse communities through the use of relevant rhetorical tools. To this end, the concept of voice operationalized in this study draws primarily on the work of Hyland (2005, 2008) and Zhao (2013) and is manifest as the use of the following rhetorical features: hedges, boosters, attitude markers, authorial self-mention, and direct reader references.

### 1.2.3 Concept-based instruction and systemic theoretical instruction

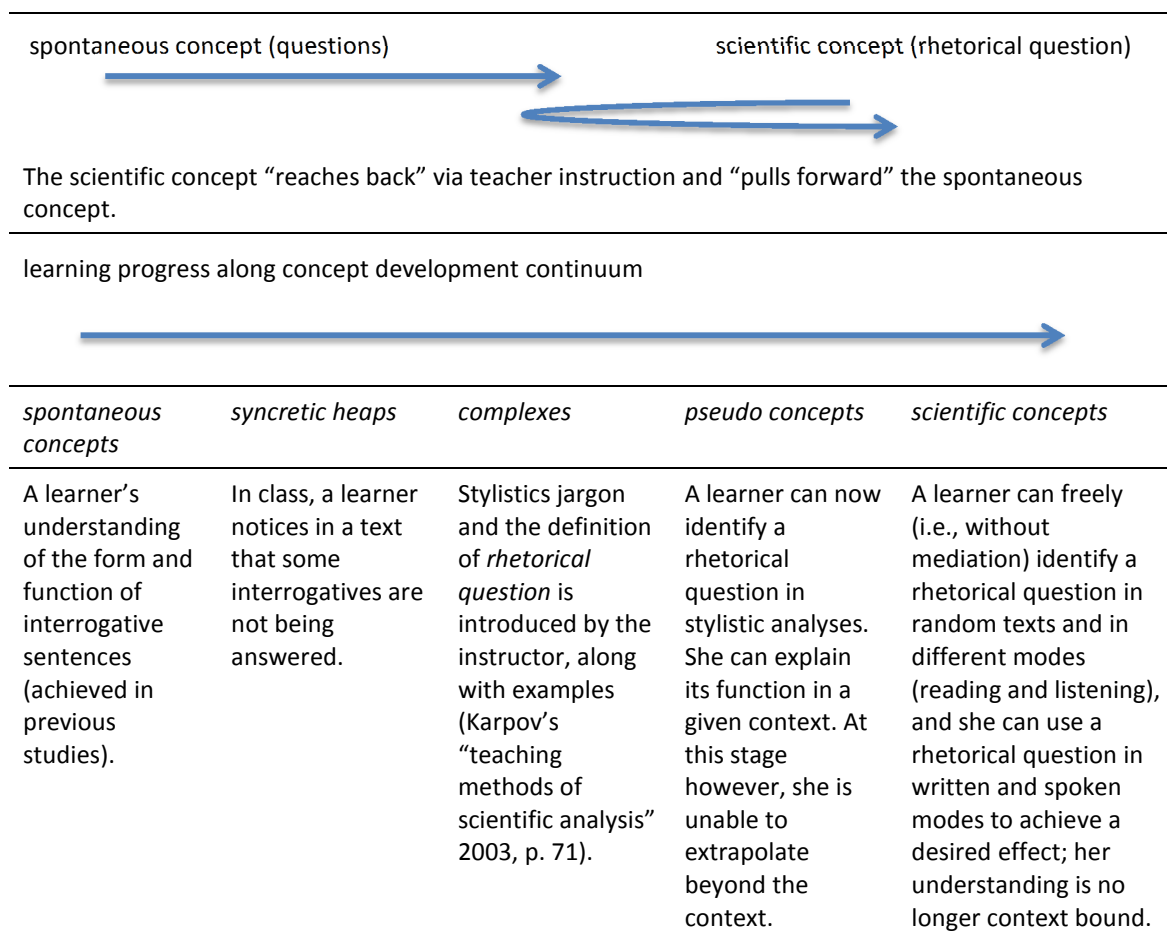
CBI operationalizes Vygotsky's theory of spontaneous and scientific concepts. Spontaneous concepts may be understood as concepts that learners acquire through everyday experiences (and are thus linked to empirical contexts) but that are not categorized or ordered. Table 1.1, using the rhetorical question as an example of a literary device (or a rhetorical feature of voice), provides a rudimentary example of how spontaneous and scientific concepts interact in the CBI context. For a detailed account of spontaneous and scientific concepts, see Chapters 5 and 6 of Vygotsky (1986), Lantolf and Poehner (2008), or Swain, Kinnear and Steinman (2010).

Although Vygotsky never developed a classroom guide for implementing scientific concept development, researchers interested in Vygotsky's theories of psychology (in particular Davydov, 1988, 1990 and Galperin, 1969, 1992) attempted to do so. This thesis study draws from Galperin's pedagogical approach to concept development.

Galperin's (1969, 1992) approach, referred to as Systemic Theoretical Instruction (STI), was devised on the basis of 800 classroom studies of different learners in varying learning contexts, including L2 instructional settings (Lantolf & Poehner, 2014). Galperin argued that the goal of STI is to move learners through stages of concept development (see Table 1.1) until they are able to internalize (or develop the scientific concept of) the target learning features. To this end, STI begins by providing learners with an explicit learning goal. Guiding development within an STI framework are material or object representations (e.g., diagrams, graphs, or pictures) of the goal. Such physical manifestations of the learning goal are paramount for Galperin because "verbal accounts are sequential and often entice learners to memorize them without deep understanding of the concept" (Lantolf & Poehner, 2014, p. 64). After an interval of study these material supports are removed and learners engage the final stages of development. The penultimate stage asks learners to communicate an action verbally to oneself (private speech) or to an interlocutor. The final phase of STI goal. Guiding development within an STI framework are material or object representations (e.g., diagrams, graphs, or pictures) of the goal. Such physical manifestations of the learning goal are paramount for Galperin because "verbal accounts are sequential and often entice learners to memorize them without deep understanding of the concept" (Lantolf & Poehner, 2014, p. 64). After an interval of study these material supports are removed and learners engage the final stages of development. The penultimate stage asks learners to communicate

Table 1.1

*Example of Concept-based Learning with Stylistics, Using the Literary Device Rhetorical Question*



an action verbally to oneself (private speech) or to an interlocutor. The final phase of STI leads to abstraction where learners can self-regulate their communicative needs. The intervention phase of this thesis study follows this program of learning. Section 4.3 of this thesis provides a more detailed overview of the stages informing STI, while Section 4.5 documents how STI is operationalized in this thesis study.

#### 1.2.4 Microgenetic analysis

Microgenetic analysis, as applied in this thesis study, aims to capture cognitive learning processes that are engaged during socially mediated interaction designed to promote learning (Lavelli, Pantoja, Hsu, Messinger, & Fogel, 2005). Microgenetic analyses have been employed successfully to track instances of learning development across various research domains, including educational psychology (Fischer & Granott, 1995; Wertsch & Hickmann, 1987; Wertsch & Stone, 1978), child psychology (Bruner, 1983; Papousek & Papousek, 1984; Trevarthen, 1977), and second language acquisition (Cheng, 2010; Knouzi et al., 2010; Park, 2012). In Vygotskian terms, the aim of a microgenetic study is to reveal cognitive uptake “in flight” (Vygotsky, 1978, p. 68) or as learning occurs. The microgenetic analysis employed in this thesis study is grounded in sociocultural theory, rooted in Vygotsky’s belief that learning is best understood by exploring its origins and the subsequent stages involved in development.

Study contexts that facilitate microgenetic research provide learners with opportunities to express cognitive processes verbally and non-verbally via interactive activities, thus allowing researchers to “track and correlate specific moment-to-moment changes to the changes observed in the participants’ cognitive performance” (Lavelli et al., 2005, p. 44). In short, this analytic lens affords researchers opportunities for scrutinizing cognitive processes (Calais, 2008). Research attempting to capture these microgenetic processes, or changes in how learners orient to target learning features, typically interprets cognitive processes along five dimensions: path (the sequence of cognitive behaviour), rate (the time and experience invested in the process of change), breadth (the transferability of the learning feature), source (the cause of the change in behaviour), and variability (how

behaviour varies across tasks and/or across learners) (Calais, 2008; Siegler, 1995, 1996).

Each of these distinct components inform cognitive uptake, and research in the field of second language acquisition (SLA) has typically examined a selection of these components in any given study. For example, Knouzi et al. (2010) traced the path, rate, and breadth of cognitive development in two French as a second language learners, while P. Chang (2010) and my earlier study (Fogal, submitted) focused on exploring only one component.

Documenting such processes informs different theories of learning and also informs pedagogy, providing insight into cognitive learning processes that practitioners can focus on when designing learning activities.

This section of Chapter 1 has highlighted the key concepts that inform this research study. In the following section I discuss my motivation for and interest in these particular areas of study.

### **1.3 Motivation and interest for this thesis study**

This thesis study developed out of two personal interests that merged early in my professional career: reading (and studying) literature and teaching academic writing. The two fused during my stay at a Japanese university, where I taught content-based classes focused on the humanities to EFL students preparing to matriculate overseas into English-medium universities. Writing assignments in these classes included personal response essays (framed as arguments) to writing prompts based on interpretations of texts from the literary canon. Of particular interest to this thesis study was a two-week module on John Steinbeck's *The Pearl*. The brief discussions I outline below—based on essays composed concerning Steinbeck's novella—emerged from informal meetings with students over two sections of the course, taught in subsequent years.

During one of these informal discussions during the first year I taught this course, I enquired about a rhetorical question employed in a student's essay. (The rhetorical question was a discourse feature that I had never witnessed this student using, nor was it a feature that I had discussed in class.) This student replied that her use of the rhetorical question was modelled on how Steinbeck employed this feature in *The Pearl*. Though interesting, I did not follow-up on her response. When I taught the same module in the following year—and concurrent with my thoughts of returning to graduate school—I had, once again, informally discussed with students writing moves associated with their essays. I enquired about two particular features on this occasion: the use of a metaphor and the use of juxtaposition (both which I had noted in discussing Steinbeck's *The Pearl*, but neither of which I suggested would be helpful for essay writing). Again, as in the previous year, these students cited Steinbeck's use of these writing features as inspiring their writing choices. The similar responses by these students—from different years I taught the course—inspired two reactions. First, their comments triggered a refreshing change in how I thought about literature use for L2 education. Although I could certainly understand (and value) the use of academic texts as models for “good writing” (I had used them successfully in my own teaching practice, and I knew from my previous studies that model essays were employed since antiquity to improve writing skills, as Smagorinsky, 1992, documented), I was also becoming increasingly frustrated by their ability to bore a class of English L2 students (this, despite my own shortcomings as an instructor). I was later relieved to learn that many colleagues had had similar experiences, and that literature use in L2, classroom-based learning contexts was a topic of debate and interest among applied linguists and educators (e.g., Gajduske & van Dommelen, 1993; Hall, 2005; Hanauer, 2001a, 2001b; Hirvela, 2001,

2004, 2005; Khati, Rezaei & Derakhshan, 2011; Paran, 2008; Vandrick, 2003). However, while employing literature was intuitively a viable alternative to using academic texts for enhancing L2 proficiency (as Hall, 2007, 2014, and Paran, 2008, emphatically noted), there remained the issue of how effective literature was at fostering L2 development (Bellard-Thomson, 2010, 2011; Carter, 2007, 2010; Edmondson, 1997; Fialho et al., 2012; Hall, 2007, 2014; Paran, 2008; Verdonk, 2002; Zyngier & Fialho, 2010). While my students' comments were initially refreshing, they were not reassuring concerning my own classroom praxis.

The second reaction that these conversations inspired had a direct consequence on this thesis study. As a result of the questions that remained about the role of literature in L2 learning contexts I began to formulate what would eventually become the underlying question informing the present research project: Does the academic writing of English language students improve as a result of studying literature? With this question in mind I scoured the Internet and the university library in Japan for answers to this question. This search resulted in my discovering pedagogical stylistics, informed in the early stages of my thinking about this project by Leech and Short's *Style in Fiction* (2007) and what I later learned to be a seminal work by Henry Widdowson, *Stylistics and the Teaching of Literature* (1975). Discovering these texts also inspired me to write "EFL Literature Studies: Student Feedback on Teaching Methodology" (Fogal, 2010), an article that examined my students' impressions about language learning with stylistics (discussed in Chapter 2 where I review recent studies on stylistics in L2 learning contexts.) However, what I did not find at this early stage was a definitive answer to my question: Does the academic writing of English language students improve as a result of studying literature?



While my professional experiences in Japan served to motivate this research project, a second major influence that informed how I would operationalize the present study transpired during my first year of doctoral studies. During this time I was introduced to a group of researchers, educators, and students who were (and continue to be) interested in the work of the Russian psychologist Lev Vygotsky and his theories of mediated learning. Mediated learning (outlined in greater detail at various points throughout this thesis study) and concept development in the tradition of Galperin (1969, 1992) emerged as a possible theoretical framework and pedagogical guide for understanding how (if at all) literature studies could help L2 learners to become better (academic) writers.

While I could merge stylistics and literature studies in L2 academic writing contexts with facets of Vygotsky's sociocultural theory to create a thesis project, I was still bereft of a manageable language construct that I could examine. Much like the interests piqued by my former students in Japan and the serendipitous introduction of sociocultural theory to my thinking, that last major obstacle to forming a research project arrived without invitation. In readings that were directed to me for other purposes, I uncovered the following: "Controlling the level of personality in a text is central to maintaining interaction with readers and building a convincing argument" (Hyland, 2004b, p. 5). This comment struck me as odd—and indeed extremely interesting. Hyland was discussing a hallmark of literary authors (their ability to control personality in a text), except Hyland was doing so in the context of academic, non-literary writing. I pursued Hyland's idea, and after further reading my thoughts converged: Through the lens of sociocultural theory I wanted to employ CBI as a framework for teaching stylistics as a tool for mediating authorial voice development. The

research questions introduced in Section 1.1 indicate how this initial plan formed a research project, while in Chapters 3 and 4 I outline how this idea was operationalized.

#### **1.4 Overview of the thesis chapters**

This thesis has six chapters. In this first chapter I briefly outlined the main foci of this study: pedagogical stylistics, authorial voice development, and CBI. I explained how these three research domains converge in one learning-to-write context. I also provided an overview of the study and introduced the guiding research questions. I then defined key concepts and introduced my personal motivation for conducting this research. In Chapter 2, I review recent work that engages the three main components of this research project and I identify gaps in the literature that this thesis study considers. Chapter 3 provides a detailed account of the research methods employed in this study. This account includes a review of how participants were recruited, a description of the instruments employed to collect data, and a description of how data were coded and analyzed. Chapter 4 outlines the classroom context that formed the intervention phase of this study. Therein I also describe my role as the instructor and researcher, and then I provide a narrative account of daily classroom interactions. In Chapter 5 I document the results of this thesis study deriving both from statistical analyses of data and from a thematic analysis of the verbal data. Finally, in Chapter 6 I summarize the findings most relevant to this study, and then I discuss the significance of the findings to relevant theoretical and pedagogical domains. I conclude Chapter 6 by outlining the limitations of this study and suggesting avenues for continued research in this and similar research contexts.

## Chapter 2

### Perspectives on Pedagogical Stylistics, Authorial Voice Development, and CBI

The purpose of this chapter is to highlight findings from prior research that are relevant to the context of this study. Specifically, this review concentrates on literature in pedagogical stylistics, authorial voice development, and CBI in L2 writing contexts. In doing so, this chapter underscores the need for and the purpose of the present study. This chapter concludes by addressing other avenues for researching the topic of this thesis and introduces the relevance of complex dynamic systems theory to interpret the findings of this study.

#### 2.1 Pedagogical stylistics

Research that has addressed the efficacy of pedagogical stylistics in L2 learning contexts is limited. While ongoing research on the pedagogy of stylistics abounds, with a focus on how best to implement stylistic analyses in the L2 (and L1) classroom, little rigorous attention has focused on the results of instructional interventions for improving students' L2 proficiency (Bellard-Thomson, 2010, 2011; Carter 2007, 2010; Edmondson, 1997; Fialho et al., 2012; Fogal, 2015; Hall, 2007, 2014; Paran, 2008; Zyngier & Fialho, 2010). Despite this limitation, and to justify the context of this present thesis study, I consider what recent publications have concluded regarding the efficacy of stylistics-based analyses for developing L2 proficiency across various learning contexts. To do so, I first address themes that are relevant to this thesis study, and then I demonstrate how many of the studies I review are not sufficiently robust to substantiate the use of stylistics for developing L2 proficiency.

Research to date has generated studies that primarily address how stylistics-based literary analyses serve to improve L2 proficiency (Allen, 2009; Paesani, 2006; Plummer & Busse, 2006; Timucin, 2001; Yañez Prieto, 2010), to enhance language awareness (Allen, 2009; Davies, 1998; Lin, 2010; Warner, 2012), and to promote analytic skills that can be applied outside a language learning context (Al-Jarf, 2007; Badran, 2012; Fogal, 2010; Plummer & Busse, 2006; Saugera, 2011; Warner, 2012; Zerkowitz, 2012). Each of these themes is reviewed below.

### 2.1.1 Pedagogical stylistics for mediating L2 performance

Allen (2009), Paesani (2006), Plummer and Busse (2006), Timucin (2001), and Yañez Prieto (2010) have reported varying degrees of learner L2 improvement as a result of stylistics-based classroom interventions. Allen (2009), in a university French foreign language (FFL) context, examined learners' self-evaluations of their writing portfolios and found that a multiple-literacies approach to FFL instruction (one that included stylistic analyses) engaged learners such that the “stylistic devices [e.g., tone and register] that they gained experience in using, provided new means of ‘playing with’ language and developing their own writerly voices [in fiction texts]” (p. 380) in the French L2. However, Allen did not expand upon this notion of “writerly voices”, nor did she define or measure development. Allen concluded that this multiple-literacies teaching approach facilitated exploring relationships between French literature and French language, resulting in improved L2 writing proficiency.

Paesani (2006) focused on a writing portfolio project in an advanced grammar and stylistics course in another university FFL course. The writing portfolio project required

students to comment on the writing progress of their portfolios and to comment critically on the writing portfolio project. Eight written pieces were included in the portfolio, and learners selected four pieces for the instructor to score. A quantitative score was assigned for process-based performance on early drafts of the four selected pieces, while a qualitative score was assigned to the final texts. Based on analyses of the portfolios, Paesani concluded that alongside improving learner's knowledge of literary texts, these learners' L2 proficiency skills (composing, revising, and rewriting) also advanced. Plummer and Busse (2006) examined how 50 EFL learners in two German universities responded to three different stylistics-based treatments. By examining term papers, a final examination, questionnaires, and video-taped focus group discussions, Plummer and Busse documented that learner writing proficiency improved.

Yañez Prieto's (2010) study focused on a literature-through-language pedagogy that was grounded in stylistics. In a literature-through-language course at an American university, 12 Spanish as a foreign language learners were first exposed to an instructional course similar to their previous learning experiences at the university, and then to an alternative stylistics and concept-based pedagogy. Yañez Prieto then compared these two methods by examining classroom activities, learning logs, interviews, and portfolios. Yañez Prieto reported that learners made L2 proficiency gains, primarily in Spanish writing: learners manipulated mediating tools (literary texts) to achieve their writing goals by, for example, playing with grammatical aspect in their essays. In contrast to studies that documented improvement in writing proficiency, Timucin (2001) reported that EFL learners in a Turkish university improved their reading comprehension skills after a stylistics-informed, language-

based intervention. The intervention employed literary texts to raise learners' awareness of the meaning-making processes informing literature.

These studies underline the promise stylistics holds as a mediating tool for developing learners' linguistic competency, with a focus on writing in L2 contexts. Below I develop the remaining two themes associated with stylistic literature, and then I discuss concerns that are present across all three of the themes addressed in this review.

### 2.1.2 Pedagogical stylistics for enhancing language awareness

The original aim of pedagogical stylistics (i.e., raising language awareness, noted by Clark & Zyngier, 2003) has been substantiated by Allen (2009), Davies (1998), Lin, (2010), and Warner (2012). For example, Davies (1998) reported on a process-based teaching method in a British university's ESL class. Davies grounded the foci of her study in Nash's (1992) work on metadiscourse and cited her own professional teaching experience to support the use of stylistics in her teaching: Davies argued that stylistics allowed learners to draw on their own experiences with the target language, thus facilitating text interpretations. Using a stylistics orientation, Davies then described a lesson plan for teaching metadiscursive functions in Ishiguro's novel, *The Remains of the Day*. She concluded that, among other benefits, a process-based teaching methodology focused on stylistics was useful for raising learners' language awareness.

Allen's (2009) study suggested that alongside developing writing proficiency, learners improved their language awareness via a multiple-literacies approach to L2 teaching that included stylistic analyses. Lin's (2010) study employed a pre-test, post-test design to examine language awareness levels in 22 EFL learners at a Taiwanese university. Statistical

analyses were used to examine a text-interpretation activity that was grounded in stylistics. Lin concluded that EFL classroom contexts that employ stylistics as a pedagogical tool can facilitate L2 language awareness competency.

Warner's (2012) study in an American, German as a foreign language university context analyzed varied stylistics-based classroom activities that informed the study of Goethe's *The Sorrows of Young Werther*. Warner concluded, based on observing classroom discussions and examining journal entries, that employing stylistics-informed pedagogies with literature studies engenders language awareness, and that such awareness facilitated "critical and creative interpretive choices" (p. 155) that serve learners in other academic contexts (the latter highlighting how pedagogical stylistics can help learners develop analytic skills, as I expound in Subsection 2.1.3).

These studies document pedagogical stylistics' ability to raise language awareness amongst L2 learners, and suggest that such awareness engenders L2 competency, and in the case of Warner (2012), improved analytic skills that are transferable across academic contexts. I examine this possibility further in Subsection 2.1.3.

### 2.1.3 Pedagogical stylistics for mediating and transferring learning and learning skills beyond language acquisition contexts

The last theme evident in recent publications focuses on claims that stylistics aids L2 learners' general academic development (Al-Jarf, 2007; Badran, 2012; Fogal, 2010; Plummer & Busse, 2006; Saugera, 2011; Warner, 2012; Zerkowitz, 2012). Al-Jarf's (2007) study, in an EFL Saudi Arabian university context, employed a treatment-then-test research design to examine how learners were able to analyze and comprehend advertisements for

stylistic features. Sixty-six learners participated in this classroom-based study. One week after a stylistics-informed treatment, and informing the post-treatment test, learners were asked to analyze an advertisement from a *Reader's Digest* magazine. Herein Al-Jarf (2007) noted that certain linguistic features were more difficult for learners to identify (e.g., learners identified more lexical features than syntactic features). Among other findings, Al-Jarf (2007) concluded that instruction in lexical, syntactic, morphological, and pragmatic features of advertisements (informed in part by stylistic analyses of advertisement text) is required in this teaching context if learners wish to better understand the language of advertising—thus informing the overall goal of their study program (i.e., a degree in translation).

Grounded in stylistics-informed, classical rhetorical analysis, Badran (2012) examined how 51 EFL students in a Lebanese university reacted to a poem. Badran examined questionnaires, writing activities, and classroom discussions. He argued that stylistics-based analyses employ “techniques available to the critic from macro-generic as well as micro-linguistic approaches to textual analysis” (Badran, 2012, p. 134), concluding that a stylistics pedagogy is a “well-rounded approach” to classical rhetoric analysis in L2 contexts.

Warner's (2012) study, cited already in Subsection 2.1.2, argued that exploring Goethe's *The Sorrows of Young Werther* through the lens of stylistics informed learners' abilities to transfer knowledge gains beyond the classroom context wherein learners are able to “pursue independent scholarship” (Warner, 2012, p. 155), thus highlighting the potential of stylistics to develop learning or analytic skills applicable beyond L2 acquisition contexts. Similarly, my own study (Fogal, 2010) and Plummer and Busse (2006) documented that classroom-based stylistic analyses in L2 contexts can also improve learners' analytic skills.



My work (Fogal, 2010), in a Japanese university EFL context, examined the preferences of 35 learners to three different teaching methods related to literary text analysis, including stylistics. In that study I examined questionnaires and interviews to identify learners' preferences, and I subsequently documented that learners exposed to stylistic analyses of literary texts improved their analytic skills. Plummer and Busse (2006) (cited in Subsection 2.1.1) also noted similar findings regarding the influence of stylistics on improving students' analytic skills.

Saugera's (2011) semester-long FFL study in an American university investigated how comparative stylistics can help intermediate- to advanced-proficiency learners examine L2 language variation. This 14-week, classroom-based exploratory case study focused on script writing and related activities as a by-product of studies in translation, translation devices, and comparative (French and English) stylistics. Based on an analysis of activities informing learners' screenplays, the screenplays themselves, and film reviews, Saugera (2011) concluded with an impressionistic account of learners, suggesting that learners were able to enjoy a "sense of linguistic and artistic accomplishment" (p. 150)—the latter suggestive of stylistics' ability to extend beyond the context of L2 acquisition by informing other (here artistic) ends.

In a Hungarian, EFL context Zerkowitz (2012) examined how Hungarian L1 students studying to become EFL teachers benefited from stylistics-based interventions during their teaching training. She concluded that stylistics-informed pedagogies helped learners make meaningful connections between language forms and functions, enabling learners to interpret texts using linguistic evidence. Moreover, drawing these connections, Zerkowitz concluded, could provide support for future EFL instructors interested in employing literature in their

classrooms, suggesting that stylistics can also inform teacher-training in diverse L2 contexts.

As these seven studies attest, stylistics-informed pedagogies have the potential to facilitate learning opportunities beyond L2 acquisition contexts. This review also documented the promise stylistics carries for developing L2 proficiency and for enhancing language awareness. These three themes underscore how L2 learners benefit from stylistics-informed literary-text analysis. However, by examining the methods by which many of the claims documented above were achieved, I concur with the concerns voiced by Hall (2005, 2007, 2014) and others regarding the validity of stylistics as a tool for developing language proficiency in L2 contexts. To better understand the cause of these concerns (and to subsequently lay the foundation for my thesis study), I now discuss limitations associated with the studies addressed above.

#### 2.1.4 Underreporting and under-collecting of data in pedagogical stylistics research

The primary concern with the studies documented above is the underreporting (or under-collecting) of data. In a research synthesis that I compiled for this thesis study (Fogal, 2015) I unpacked this concern and I suggested strategies for improving the research agenda of pedagogical stylistics. As I noted in this forthcoming publication, this concern is bound by four themes: (a) underreporting facilitating conclusions based on intuition and trust; (b) underreporting in quantitatively oriented studies; (c) underreporting of important data; and (d) underreporting of different L2 contexts. I expand upon each theme below.

Hall's (2005, 2007, 2014) concerns that the validity of L2 stylistics is based on trust (2005) and intuition (2007) is substantiated by many of the studies reviewed above (Allen,

2009; Paesani, 2006; Plummer & Busse, 2006; Saugera, 2011; Yañez Prieto, 2010). Below I briefly outline reasons for these concerns.

First, Allen's (2009) claim that learners were able to improve their "writerly voices" is undocumented. Readers of Allen's study were not informed how voice was defined, coded, or analysed. Second, Saugera (2011) relied on personal experience to draw conclusions about learners' linguistic (and artistic) accomplishments with the study material, but data regarding how accomplishments were measured or collected is not documented. Instead, readers are left to trust how Saugera intuited learning outcomes. Third, Plummer and Busse's (2006) study suggested that learners' writing proficiency improved after a stylistics-informed study program. However, the study did not report data that could substantiate this change. Similarly, despite collecting relevant data, neither Paesani (2006) nor Yañez Prieto (2010) reported evidence that could support their findings. Paesani (2006) documented the quantitative measures taken to collect data but failed to record the resultant data: readers learned about coding procedures and scoring techniques, but the scores were left unreported. Instead, readers were told simply that studying literature and language in tandem "pushed them [learners] toward more advanced competencies in the second language" (Paesani, 2006, p. 628). Similarly, Yañez Prieto (2010) described how participants in her study could "recontextualize mediating tools" (2010, p. 69) that resulted in L2 writing development, but she did not document how, for example, improvement was measured. In each of these cases insufficient data were reported. Accordingly, readers are unable to reach informed conclusions about the findings and are left, instead, to trust how the respective authors interpreted the data.

The next concern regarding underreporting involves the degree of statistical analyses reported in quantitatively oriented studies. For example, no studies reported any measures of central tendency. Plummer and Busse (2006) noted that a small sample size precluded subjecting their data to statistical analyses; other studies did not justify the limited reporting of data. Moreover, Paesani (2006) and Yañez Prieto (2010) chose not to report empirical data that was collected, and only Timucin (2001) employed a control group. No studies reported effect sizes. The lack of data reported in these studies precludes measures of overall effect size (Norris & Ortega, 2000, 2006), further limiting the possibility for a meta-analysis of this research domain (Cumming, 2012; Ellis, 2010).

Underreporting relevant data is also a concern in L2 stylistics-oriented research. The limitations here constrain how these studies can be interpreted and also restrict replication studies. For example, Allen (2009), Badran (2012), Davies (1998), Paesani (2006), Timucin (2001), Warner (2012), and Zerkowitz (2012) chose to omit details regarding their research design. Moreover, many of the observations documented in these studies did not indicate what instruments were employed to collect data, nor did these studies indicate how data were coded or analyzed. Further, many published studies did not report context-related data: Allen (2009) and Paesani (2006) did not report the countries where their studies occurred, while Allen (2009), Davies (1998), and Warner (2012) excluded mentioning the number of participants in their studies. The L2 proficiency levels of participants were also frequently undocumented (e.g., in Allen, 2009; Al-Jarf, 2007; Badran, 2012; Lin, 2010; Timucin, 2001; Warner, 2012). Research to date in pedagogical stylistics has also been underrepresented in regards to different research contexts. For example, studies have largely concentrated on

foreign language contexts, on university-aged learners, and on classroom-based research, but not other populations or educational settings.

Future research into pedagogical stylistics would benefit from addressing many of the shortcomings listed above. For example, expanding the research focus would develop a more robust collection of studies in this research domain; documenting more data could strengthen claims and within the context of this research could help validate the effectiveness of stylistics as a tool for improving students' L2 proficiency. Reporting more data could also help inform subsequent research and would also afford readers the opportunity to interpret the importance and relevance of findings to their own educational contexts.

While the present review has highlighted legitimate concerns, this critique should not be read as an attempt to undermine the findings I have discussed. Rather, I have attempted to demonstrate potential weaknesses in how this study domain has historically collected and reported on data, and I have attempted to substantiate the concerns of Hall (2005, 2007, 2014) and others regarding research quality on pedagogical stylistics. In doing so, this review underscores the need for a more robust study agenda in this context that includes wider research foci and more attentive reporting apparatuses. This review also provides multiple avenues for further studies that can bridge the gap between the concerns of stylistics and those of L2 teaching and learning, as Paran (2008) and others have noted. To address these issues, the design of this thesis study has attempted to account for some of the concerns outlined above.

In the next section of this chapter I discuss recent studies that have examined voice construction in L2 writing contexts.

## 2.2 Authorial voice development

As outlined in Chapter 1, voice is defined in this study as an amendable, context-specific representation of self that is manifest through lexical, syntactic, and rhetorical choices woven together as result of the linguistic interplay between the self and the social. Although there is presently no single, unifying theory of second language writing (Cumming, 2010, 2013a; Grabe, 2001) through which to examine the construct of voice—Cumming (2013a) argues that the absence of a unifying theory is a result of “too many contradictory purposes, situations, and conceptual issues that it would have to serve” (p. 1)—the absence of a single theory does not preclude employing a conceptual framework for understanding and explaining L2 writing acquisition (Cumming, 2010; Silva & Matsuda, 2010) or, in the context of this thesis study, voice construction. The analysis of voice I present in this thesis study is grounded in sociocultural attempts to examine microgenetic processes associated with L2 authorial voice development. To justify this microgenetic approach to examining voice, I next review recent publications that have addressed present understandings of authorial voice development in L2 writing contexts.

By engaging authorial voice as a socially-constructed, dialogic representation of self, research in this domain has tended to take up two particular themes: macrolevel studies that investigate how large-scale phenomena, such as learners’ socio-cultural histories and their negotiations with disciplinary discourses influence voice development, and microlevel (i.e., microgenetic) studies that investigate learners’ cognitive processes that facilitate voice development. I review each of these themes below.

### 2.2.1 Macrolevel research attending to voice development

As I remarked in Section 2.2, macrolevel studies of voice development approach development using a wide analytic lens. These studies have examined development through “the person behind the written words” (Hirvela & Belcher, 2001, p. 85), highlighting the constraints and affordances with which L2 learners engage in the larger context of their personal histories and their discourse communities. For example, Hirvela and Belcher challenged both the notion that L2 writers must be taught how to develop their authorial voice and the underlying premise that L2 writers are, then, voiceless in their L1. They identified the L1 writing histories of three mature, English L2 graduate students in the United States and examined if and how their L1 voices constrained or assisted their L2 authorial voice development. Hirvela and Belcher concluded that constructing an L2 voice initiates in the L1, highlighting the influence of macrolevel instances of language use that contribute to authorial voice development. Matsuda (2001) noted similar findings regarding Japanese L1 (while also underlining that constructing voice is not a universal process), while Shen (1989) described his own similar process of acculturation from his L1 Chinese as he negotiated his English L2 writing development, including his authorial voice. In a similar vein, Bayyurt (2010) examined voice construction (operationalized as hedges and boosters) across Turkish and English language essays composed by 74 freshmen, undergraduate L1 Turkish students in an English language teaching program across two Turkish universities. The study compared 74 essays written in Turkish with 74 essays written in English. Both sets of essays comprised responses to the same essay prompt. Bayyurt found that writing practices in students’ L1 influenced how learners constructed their voices in L2 English academic writing. Similar findings were reported by Can (2006), again in the Turkish L1 context, and

by Wu and Rubin (2000) who examined how English and L1 Chinese interacted (see Fløttum, 2012, for an extensive discussion of how L1 writing practices influence L2 voice construction). These studies suggest that authorial voice development is a process of acculturation that initiates in learners' personal histories, and that the process of development is embedded in a social activity guided by language: Writers engage with their histories (and as I demonstrate below, their discourse communities) to adopt and adapt, and to acculturate through and with, a developmental process on a macro scale.

Cumming (2013b), in discussing the findings detailed in Kim et al. (2006), highlighted struggles encountered by L2 learners as they acculturated to new discourse spaces. Cumming outlined the obstacles informing voice development over three years of study in two Japanese ESL learners: Rihoko, initially an architecture student, and Kazuko, a political science student. The obstacles that Cumming described were heavily socio-cultural, informed by students' (a) initial impressions of how to employ personal expressions of authorial self (derived in part from a one-year ESL course they took prior to matriculating) and (b) expectations about their writing once they matriculated, informed by their programs of study. Cumming emphasized the social practices informing voice construction in these two learners. For example, Kim et al. noted that Rihoko had considerable problems acculturating to the writing demands of her academic community. She was required to invoke a more personal voice in her writing than she was accustomed to or that she believed was necessary, with the resulting struggles informing her decision to change fields of study. In contrast, while Kazuko was eager to promote an engaging voice in her writing—one that was nurtured during her yearlong ESL studies—assessors of her work asked that she produce a more nuanced self in her writing. Kazuko too experienced struggles as she embraced this



mode of self-presentation (although she did not change her field of study). These experiences, Cumming suggested, revealed the complexity involved in, for one, developing an authorial voice as L2 learners engage with new, academic discourse spaces. As Kim et al. discussed, previously-held beliefs become challenged and lead to struggles that have important and long lasting consequences. Cumming concluded that educational practices (informed by further studies in this area) should facilitate opportunities for multilingual students to negotiate, among other considerations, “expressions of personal identity [including authorial voice] in relation to relevant discourse communities” (p. 145). Cumming and Kim et al., then, discussed the complexity of constructing voice, a process that is informed by learners’ personal histories and their immediate negotiations with new discourse communities.

Similar findings have been reported in research publications across varying contexts. For example, alongside discourse-based interviews with 23 students, Hyland (2012) compared the use of questions, reader references, directives, attitude markers, hedges, and boosters (as constituents of authorial voice) across a corpus of 64 Hong Kong undergraduate students’ reports with a reference corpus amassed from published research articles. Hyland found that undergraduate students’ uses of these features (with an emphasis on attitude markers, reader references, hedges, and boosters) differed considerably from how authors of published research articles employed the same features. Hyland documented that undergraduate writers in this context were conscious of the need to balance demonstrating their knowledge with deference to their readers (here, professors assessing their final year projects), and that such awareness often forced these writers to foreground their identities to the confines of their disciplinary conventions at the expense of their personal voices. In short,

writers' immediate contexts (attempting to adhere to the writing conventions of their academic discipline) both informed their voice development, but also constrained such development, suggesting, as Cumming (2013b) and Kim et al. (2006) suggested, that constructing authorial voice is simultaneously informed by one's beliefs about writing practices and the immediate writing contexts in which learners are engaged.

While Hyland's (2012) work examined writing practices across disciplines, Guinda (2012), like Kim et al. (2006), provided a more focused examination of a particular discourse community. Guinda examined the role of voice (and stance) in the practice of communicating about visual data (graphs, diagrams, and tables) amongst students in an aeronautical engineering program. Employing a mixed-method research design (discourse-based interviews and discourse analysis) Guinda investigated the writing practices of second-year students enrolled in her academic English class at a Spanish university. She examined the stance and engagement features (collectively described as voice) in 475 commentaries explicating a variety of diagrams (a pie chart, a tree diagram, a ring graph, a flow chart, and two line graphs). While Guinda's students favoured employing engagement features such as interactive discourse markers (directives and reader pronouns) in their writing, this tendency was a result of the "indiscriminate imitation of [writing] features" (p. 169) that failed to completely address the needs of the discourse. This process resulted in what Guinda described as "frustrated proximity": Although students' authorial voices were formed primarily by their disciplinary expectations, these expectations were misconstrued (i.e., students often failed to address the appropriate discourse communities in their descriptions). Guinda advocated for a socioliteracy dimension to students' study programs that could help writers "discern what communities they need to address and through what linguistic and

rhetorical features” (pp. 180-181). While emphasizing the role of learner engagement practices on influencing voice development within a focused discourse community (here, writing in an aeronautics engineering program), Guinda’s study also corroborated the aforementioned findings that examined how learners interact with discourse spaces. Similar findings across other contexts have also appeared. For example, Bloch (2007), Coffin and Hewings (2005), Lam (2000), Y. Chang (2010), and Warschauer (2002) all provided further examples of voice development across different macrolevel contexts that were mediated by online writing communities or online writing tasks. These studies highlight online interaction as another instance of social practices that facilitates voice development in writing.

While the studies described thus far focused on macrolevel influences on voice development, Ivanič and Camps (2001), while incorporating learners’ histories into their study, also focused on the struggles learners encountered while negotiating their identities at the microlevel—though not microgenetic level—(manifest as an amalgam of “voice types”) in new discourse communities. To do so, Ivanič and Camps examined the lexical, syntactic, and rhetorical choices in the writing of six Mexican graduate students studying in the United Kingdom. They also conducted interviews with these learners and with readers that assessed their course-based writing tasks. Employing Halliday’s (1985) macro-functions of language, Ivanič and Camps suggested that these writers adapted three “voice types” (p. 10) related to their academic identity that existed jointly throughout their writing: ideational positioning (writing on a topic or expressing a particular view), interpersonal positioning (writing that engages the reader-writer dialectic by displaying the author’s authority—or lack thereof) and textual positioning (shaping the text, or how authors position themselves relative to the genre of the writing task). Ivanič and Camps noted that voice types informing identity were a

socially constructed amalgam (in the Bhaktinian sense—see Wertsch, 1991) that writers negotiated as they acculturated to their academic disciplines, and that writers' identities (and the resultant voice) were manifest in how learners positioned themselves through their writing. Ivanič and Camps, then, were able to account for macrolevel and microlevel (though not microgenetic) developmental features of authorial voice.

With the exception of Ivanič and Camps (2001), I categorized the studies discussed thus far as exclusively macrolevel studies. This description is accurate insofar as these studies analyzed authorial voice development through a wide analytic lens that addressed learners' personal histories or that situated learners in discourse-based contexts, or both. As I remarked, these studies emphasized the influence these extensive communities have on voice development. However, absent from these studies—though through no fault of the authors, given their respective foci—are reports on microlevel activities that reveal the cognitive processes informing voice development in individual learners. I now turn my attention to this perspective.

### 2.2.2 Microlevel research attending to voice development

Numerous studies have examined the development of L2 writing processes at the cognitive, microlevel (see Cumming, 2001, Leki, Cumming & Silva, 2008, related chapters in Manchón, 2009, the research program of Manchón and colleagues, described in Manchón, Roca de Larios & Murphy, 2009, and studies through the lens of complex, dynamic systems theory, such as Baba & Nitta, 2011, 2014, De Angelis & Jessner, 2012, Larsen-Freeman, 2006, Spoelman & Verspoor, 2010, and Verspoor, Schmid, & Xu, 2012), but few studies have attempted to provide a microgenetic account of a specific writing construct like

authorial voice. I discuss these limited studies on voice here. The two studies described below—one of which I conducted and which is based on the pilot phase of this thesis study (Fogal, submitted)—explicitly posited that voice construction occurs within a dialogic space. Accordingly, these studies take as given the macrolevel factors noted above that influence voice, and they focus, rather, on exploring microgenetic cognitive processes informing voice development. While P. Chang’s (2010) study investigated the cognitive processes implicated while learners constructed voice in their writing, my study (Fogal, submitted) focused primarily on the sequence of cognitive activity. Together these studies provide an initial glance at the microgenetic processes informing voice development and provide detailed accounts of cognitive behaviour in this regard. I discuss each of these studies below.

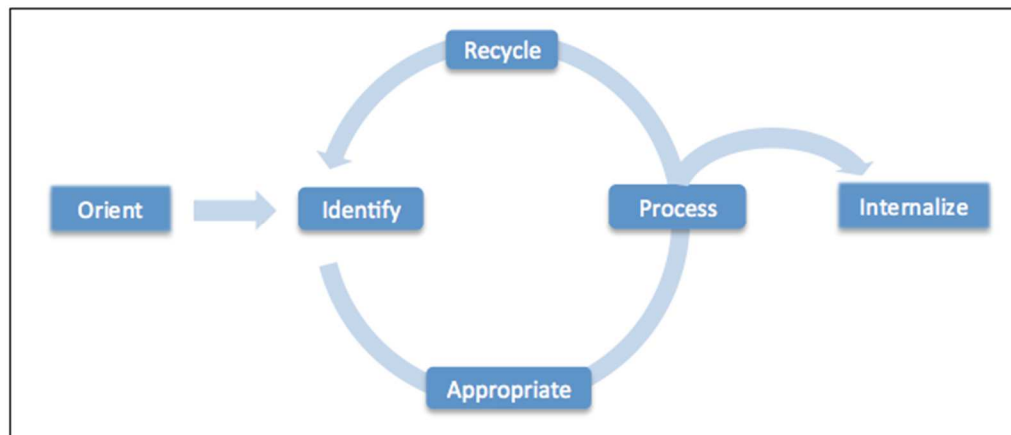
P. Chang (2010) examined how seven Chinese ESL post-graduate students in the United States engaged with a linguistic resource tool (Chang refers to this tool as a stance corpus) for supporting voice development. The stance corpus was drawn from introductory passages of published research articles in education, and learners engaged the corpus across three, one-hour learning sessions. Chang documented a positive correlation (statistics were unreported) between stance markers that learners’ employed and their improvement in academic writing (specifically, the writing of introductions). Relevant to the present thesis study, and among other foci, Chang investigated the cognitive activities that learners engaged in while interacting with the stance corpus. To capture these cognitive processes Chang employed stimulated recall interviews and text analyses of one pre-task writing assignment and of three drafts of a paper from each learner. By employing a taxonomy of thinking skills describing how learners engage with corpora (O’Sullivan, 2007), Chang documented that the stance corpus elicited higher-order thinking, noting that reasoning skills, exploring skills, and

inferential skills were the most dominant interactive cognitive processes that informed voice development. These findings provide initial insight into the processes implicated during corpus-mediated voice construction. These findings, however, should be interpreted with caution. For one, the findings are constrained by the length of the intervention. While three hours is sufficient for initially tracking change over time—especially when examining a novel construct (from a microgenetic perspective) like voice—a more longitudinal study of voice development would likely yield more sound results. Second, while Chang’s study was able to reveal cognitive process involved in voice development, her study was unable to analyze the sequence of these events (nor was she able to uncover other dimensions of microgenesis): Understanding the sequence of cognitive events would contribute a more robust framework for considering theoretical and pedagogical implications.

Despite these shortcomings, P. Chang (2010) initiated the early stages of mapping the cognitive processes involved in constructing authorial voice in writing. While limited to a particular context framed by corpora-initiated learning, Chang’s findings are nonetheless of interest. These preliminary findings were the first to document cognitive behaviour regarding voice development, and they were the first to record the type of thinking that learners engage in while constructing authorial voice. Moreover, from a microgenetic perspective, Chang’s work initiated studies on a previously unexamined construct and has subsequently invited further research on this phenomenon. The study I conducted on the pilot phase of this thesis study (Fogal, submitted) attempted to better understand voice development at this microgenetic level.

My earlier study (Fogal, submitted) focused on cognitive changes that language learners experienced as they developed their authorial voices. By employing literary texts as

models for constructing voice in essay writing, I investigated microgenetic changes in the development of two constituents of authorial voice: boosters and direct reader references. I followed two university-aged Japanese ESL learners preparing for a high-stakes writing test, and employed a pretest-posttest research design to trace development over an eight-day intervention period. Development was tracked by examining five essays that the two learners each composed, audio-recordings of in-class discussions between the two learners, and notes that I collected as a researcher during the intervention period. The preliminary findings revealed that learners engaged a developmental sequence of cognitive processing characterized by the following stages: orient, identify, appropriate, process, recycle, and internalize. In Fogal (submitted), and reproduced in Figure 2.1, I demonstrated how these stages triggered a cyclical pattern of cognitive uptake.



*Figure 2.1.* Cyclical pattern of cognitive learning processes for authorial voice development.

The first stage constituted learners orienting themselves to the target language by means of explicit teacher instruction. The second, third, and fourth stages involved learners identifying the target language in use, appropriating the target language, and then processing feedback on their use of the target language, respectively. Learners in this previous study

then moved to stage five where they would begin anew to cycle through (or recycle) through stages two, three, and four with the aim of internalizing the target feature. While the participants exhibited unique learning trajectories, both participants required numerous passes through this cycle as they attempted to internalize the target features. While the findings reported in Fogal (submitted) were consistent with usage-based, emergentism theories of SLA (Eckstein, Chariton, & McCollum, 2011; Hegelheimer, 2006; Hegelheimer & Fisher, 2006; Larsen-Freeman, 2011, 2013), documenting these stages (or cycle) of development was also consistent with Siegler, (2006) and Siegler and Svetina (2002), who argued that general patterns of learning development may be present despite individual variability (see also, Larsen-Freeman, 2011). Not without its limitations—for example, my previous study did not account for individual variability which could have shed light on individual developmental processes (see Verspoor et al., 2012, for a study that focused on individual variability in L2 writing)—this previous study of mine confirmed earlier research through a sociocultural theory lens (Cheng, 2010; Ferreira, 2005; Ferreira & Lantolf, 2008; Park, 2012) that demonstrated a nonlinear path of cognitive development in L2 writing acquisition while, moreover, further substantiating the work of Cheng, and Park, in other learning-to-write contexts, that documented a cyclical, cognitive learning process regarding the uptake of L2 writing. Together, the studies of P. Chang (2010) and Fogal (submitted) provided a detailed account of the cognitive processes informing voice development, thus expanding understandings of how learners negotiate the complex task of constructing voice.

This review has highlighted how current conceptualizations of voice development are rooted in dialogic and socio-cultural contexts. For example, Bayyurt (2010), Hirvela and Belcher (2001), Ivanič and Camps (2001), and Kim et al. (2006) all demonstrated that writers



engaged both with their histories and with their discourse communities as they negotiated the complex process of constructing authorial voice. Furthermore, this review has also documented microgenetic, cognitive engagement features (types of higher-order thinking) triggered during voice development (P. Chang, 2010) and the sequence of cognitive processing in voice development (Fogal, submitted) across two distinct learning contexts. However, the latter microgenetic studies remain to be substantiated. Accordingly, one aim of this study is to investigate the five dimensions of microgenetic change (path, rate, breadth, source, and variability, as documented by Calais, 2008 and Siegler, 1995, 1996) concerning voice construction. In doing so this study aims to widen microgenetic understandings of voice development in English L2 learners.

In the first two sections of this chapter I outlined the need for research to address important concerns in the areas of pedagogical stylistics and voice development. In the subsequent section of this chapter I discuss the contributions of CBI to L2 academic writing contexts, and then I describe how this thesis study aimed to further advance this domain.

### **2.3 Concept-based instruction and systemic theoretical instruction**

Numerous studies have investigated CBI in various L2 contexts (Arievitch & Haenen, 2005; Brooks et al., 2010; Gánem-Gutiérrez & Harun, 2011; Garcia, 2012; Kim, 2013; Knouzi et al., 2010; Lai, 2012; Lapkin et al., 2008; Lee, 2012; Negueruela, 2003; Swain et al., 2009; Twyman et al., 2003; van Compernelle, 2011, 2012), but only limited studies have examined CBI in L2 academic writing contexts (Ferreira, 2005; Ferreira & Lantolf, 2008; Johnson, 2008; Yañez-Prieto, 2008, 2010). These latter studies all tracked writing development qualitatively and quantitatively in L2 courses over a university semester. I

address these studies here.

Alongside their writing development, Ferreira (2005) tracked the theoretical thinking and the meaning-making processes of L2 English learners in a genre-based writing course. Ferreira's study traced the genre development of announcements, cover letters, and argumentative essays across various tasks. The data were collected from 14 ESL students in a freshmen composition course at an American university. Ferreira's findings were: (a) some students (six) still resorted to empirical, rather than theoretical thinking; (b) writing improved in some areas (cover letters), but not across genre-types (argumentative texts); (c) most students did improve in genre areas, but most still maintained the five-paragraph format; and (d) attempts at concept-development were "highly affected" by previous educational experiences (p. iv). Despite these findings, data suggested that students were engaged in theoretical thinking to varying degrees, and that this thinking informed student writing in different contexts. Regardless, Ferreira reported marginal evidence for CBI practices in genre-based writing contexts, and she called for continued studies to examine the potential of CBI as a mediating tool in L2 writing contexts.

In the same genre-based context that followed the same 14 ESL learners as Ferreira (2005), Ferreira and Lantolf (2008) investigated learners who were practicing their writing by focusing on three writing tasks: short answers, long pieces (1 page, or more), and transformed practice, where learners transformed their writing into another genre. Ferreira and Lantolf's findings echo those of Ferreira (2005): (a) learners experienced difficulty orientating to a theoretical-based teaching/learning style; (b) writing in the genres in question improved, but Ferreira and Lantolf were unable to determine if this was due to the CBI intervention, or if learners had resorted to their traditional (empirically-based)

understandings; and (c) learners were able to develop a conceptual understanding of genre, but learners were sometimes unable to employ this tool to develop their writing. Again similar to Ferreira's findings, learners were able to demonstrate the scientific-concept development of genre, but this too was inconsistent. These inconsistencies across learner groups may be explained by the CBI approach employed by Ferreira, and Ferreira and Lantolf. Although genre-based pedagogies consider writing development to be a by-product of instructor intervention (Ferreira & Lantolf, 2008), both of these studies employed a learner-driven approach to concept development designed by Davydov (1988, 1990). A CBI approach using Galperin's (1969, 1992) STI (which is more instructor driven, Ferreira, 2005) may have produced alternative findings. The findings of Ferreira, and Ferreira and Lantolf are reminiscent of the criticisms discussed above regarding pedagogical stylistics: findings support only the potential benefit of CBI interventions, specifically in genre-based writing contexts. Other studies, however, within similar (Johnson, 2008) and wider L2 writing contexts (Yañez-Prieto, 2008, 2010) have demonstrated more meaningful results with CBI-informed teaching methods.

As research-instructor, Johnson (2008) employed a CBI-informed pedagogy to develop learners' theoretical understanding of genre and to understand how learners apply this understanding across different writing tasks. As part of an action-research study, Johnson followed the work of three freshmen ESL academic writing students enrolled in his composition class at an American university. His study extended over one, 16-week semester. Johnson demonstrated that learners internalized genre-based scientific concepts via CBI-informed mediation. Unlike Ferreira (2005) and Ferreira and Lantolf (2008), Johnson applied Galperin's (1969, 1992) approach to CBI. (This difference is particularly relevant for

this present thesis study as I adopt Galperin’s theories of educational development to my intervention design—I elaborate upon Galperin’s STI and the design of my intervention in Sections 4.3 and 4.4.) Johnson’s findings are notable for two reasons. First, Johnson’s work emphasizes the potential of CBI for mediating L2 writing development. For example, learners in Johnson’s study took advantage of a Schema of a Complete Orienting Basis of an Action—a primary mediating tool of CBI—that initiated students’ internalizing of genre as a concept and that initiated the subsequent control of the concept to address their communicative (writing) needs. Second, Johnson addressed the efficacy of employing Galperin’s approach to CBI in writing contexts (as opposed to that of Davydov’s, 1988, 1990)—an approach recently supported in the literature as the favoured interpretation of Vygotsky’s theory of spontaneous and scientific concepts in L2 classroom-based learning contexts (Lantolf & Poehner, 2014). Johnson’s findings provide initial support for CBI-informed practices (in the tradition of Galperin) in L2 writing contexts, though further studies are needed to corroborate his work.

While Ferreira (2005), Ferreira and Lantolf (2008), and Johnson (2008) all tracked writing development in genre-based writing contexts, Yañez-Prieto (2008) tracked writing development (among other foci) in a literature-through-language course. Yañez-Prieto’s study (outlined in Subsection 2.1.1) was conducted with an L2 Spanish class at an American university. In a role similar to Johnson (2008) as researcher-instructor, Yañez-Prieto focused on mediating learners’ understanding of how literary authors manipulate language (e.g., genre, metaphor, grammatical aspect, or punctuation) with the objective of facilitating L2 language use that could evoke similar (emotional) responses from readers or interlocutors. More specifically, one aim of Yañez-Prieto was to examine the influence of a CBI-informed

pedagogy (in the tradition of Galperin, 1969, 1992) on developing L2 Spanish writing proficiency. As I documented in Subsection 2.1.1, Yañez-Prieto found that CBI had a positive impact on writing (with an emphasis on grammatical aspect) that resulted from recontextualizing concepts (grammatical aspect) mediated by literary texts. More than just corroborating other findings on the efficacy of CBI-informed classroom teaching that can lead to improved L2 writing proficiency, Yañez-Prieto's research also addressed how literature studies interact positively with a CBI-informed pedagogy to improve L2 proficiency.

While discussing the purposes of L2 pedagogy Negueruela and Lantolf (2005) noted that one goal of instructed L2 acquisition is to empower language learners to “exercise their agency over the language when communicating... (and create) a deep understanding of the possibilities offered them by the language to construct the appropriate meanings the user wishes to construct” (p. 7). Collectively, the research to date has demonstrated the potential of CBI to mediate this pedagogical end in the L2 writing context. Across varying degrees and contexts learners internalized target language features such as genre and grammatical aspect to meet their (linguistic) expressive needs. However, more studies that examine the role of CBI are needed to better understand its benefits for L2 learners across different L2 writing contexts, including authorial voice development.

In the previous three sections of this chapter I outlined the need for research to address important issues in pedagogical stylistics, voice development, and CBI. In discussing these areas I highlighted how each research domain can uniquely benefit from this thesis study. Accordingly, the aim of this study may be interpreted as multidimensional, and potentially lacking focus as a result. However, these multiple domains contribute to a novel

learning-to-write amalgam that facilitates understanding of each component. While each research domain can uniquely benefit from this study, I suggest that these domains also operate together to inform findings on the individual research foci comprising this thesis project.

In the next section of this chapter I discuss other potential research approaches that could address the concerns outlined above.

#### **2.4 Other research domains**

Some of the themes examined in this study could benefit from varied approaches to L2 teaching and learning. For example, a systemic functional linguistics (SFL) approach to literary text analysis could reveal how written texts are an amalgam of choices that reflect the constraints or freedoms of a given social context. Accordingly, exposure to how texts generate meaning may help L2 learners make informed decisions about their communicative needs (Herriman, 2013) and may help learners develop their L2 proficiency (see Coffin & Donohue, 2012, for an extensive review). However, as an objective of the present thesis was to help the participants improve their academic writing and their authorial voice, my own limited experience with SFL as a tool for mediating classroom-based L2 acquisition would have increased the likelihood that participants were being limited in their writing development by my own limited knowledge of SFL and therefore shortcomings as an instructor.

Another possible avenue for pursuing this research topic concerns the creative form (i.e., artistic expressions manifest as literary texts) as a starting point for mediating L2 academic writing development—as Allison (2004), Hanauer (2011), and Lobman (2010)

have called for. However, employing this perspective in this thesis study would have overlapped greatly with previous studies (Belcher & Hirvela, 2000; Hanauer, 2010; Hirvela 2001, 2004, 2005; Iida, 2012; Marsen, 2012; Spiro, 2010; Vandrick, 2003; Wang, 2009) that already have convincingly demonstrated the benefit of classroom-based interactions that focus on creative forms for L2 writing development (e.g., *haiku* poetry, as in Iida, 2012, or poetry studies in general, e.g., Hanauer, 2011). Moreover, as Bradford (1997) and Carter and Nash (1990) have emphasized, stylistic analyses need not be reserved for literary (or creative) texts, and approaching this study as learners interpreting creative acts would misrepresent stylistics as an analytic tool for exploring literature rather than a tool for exploring written discourse in general.

This research study could also have benefited from an analysis of literary texts and participants' essays through the lens of contrastive rhetoric (as originally conceived by Kaplan, 1966, and then reconceptualised as intercultural rhetoric by Connor, 2004, 2008). Connor emphasized the study of genre-types with a particular focus on processes and contexts that inform the writing process. The pedagogical aim of these foci is to help L2 learners in uncovering these practices, with the intent that awareness will aid L2 development. While approaching my work through this lens could be interesting—particularly to analyze how L2 learners' knowledge of the histories and perspectives of poets and fiction writers contribute to creating literary texts, and how this knowledge could improve L2 writing development—this approach would have shifted the focus in my study from directly aiming to address participants' writing goals to foregrounding the processes, contexts, and particular situations informing literary texts and contrasting those in the L1 and

L2 (rather than foregrounding the texts themselves, a distinction discussed in Subsection 1.2.1).

One final and relevant perspective for viewing this thesis study is complex dynamic systems theory applied to SLA (here I include complexity theory and emergentism, following Dörnyei, 2014, who classified these theories as interrelated approaches to understanding the behaviour of dynamic systems). While this approach did not inform the original design of the present study, its relevance emerged later while interpreting the findings. Accordingly, while I chose not to interpret this study through the lens of SFL, understanding the creative form, or intercultural rhetoric, the (seemingly) compatible views of dynamic systems theory with sociocultural theory—the primary theoretical lens informing this thesis study—merit the inclusion of complex dynamic systems theory in this thesis study. For example, both theoretical approaches to SLA see the learner and the learning context as inseparable, and as such both theories try to unify the social and the cognitive dimensions of L2 acquisition. Both theories also argue that classroom-based language use affects, via self-organization and emergence, how language develops, and both theories see language learning as non-linear and as emerging from individuals interacting (Larsen-Freeman & Cameron, 2008). More specifically, for example, the Vygotskian notion of spontaneous and scientific concepts (that I introduced in Subsection 1.2.3) can be interpreted as different attractor states in the parlance of complex dynamic systems theory: as learners move from one concept to another as they develop their L2 proficiency, so too do learners shift from one attractor state into another as their L2 proficiency develops. These and other points of compatibility between sociocultural theory and complex dynamic systems theory



are further supported by, and elaborated on, in de Bot, Lowie, Thorne, and Verspoor (2013), Ellis (2011), Jörg (2009), and van Geert (1998).

To be sure, these theories are not wholly congruent. For example, sociocultural theory posits that learner development originates in the external (i.e., in socially and culturally situated contexts), while complex dynamic systems theory argues that learner development originates from interaction between internal (i.e., cognitive) and external factors and contexts (Larsen-Freeman & Cameron, 2008). However, such inconsistencies do not preclude compatibility. Rather, objections that the two theories may be discordant arise from, for example, the position that learners' volition in L2 learning is not compatible with complex systems that "spontaneously self-organize" (Larsen-Freeman, 2006, p. 613)—see Kirshner and Kellogg (2009) for a thorough discussion of this incongruity. Despite these objections to their compatibility, I maintain, alongside Larsen-Freeman, Ellis, and others that sociocultural theory and complex dynamic systems theory are compatible and complementary theories of SLA (this compatibility is fleshed out further in Chapter 6 where I discuss the findings from this thesis project). Accordingly, I decided that reviewing literature relevant to L2 writing development and complex dynamic systems theory, in the subsequent paragraph, was pertinent for this study.

To date, present investigations into different components of L2 writing development through the lens of complex dynamic systems theory have included, but are not limited to, examinations of fluency (Baba & Nitta, 2011, 2014; Larsen-Freeman, 2006) and relatedly, accuracy and complexity (Baba & Nitta, 2011; Larsen-Freeman, 2006; Spoelman & Verspoor, 2010), grammatical metaphor (Byrnes, 2009), and assessment (De Angelis & Jessner, 2012; Verspoor et al., 2012). While having addressed like and distinct components

of L2 writing, findings from these studies reported similar developmental processes. For example, all of these studies noted that writing development across learners was characterized by either non-linear, erratic, or emergent behaviour. Emergent behaviour in dynamic systems occurs when there is a strong shift in learners' orientation to the target language, resulting in learners reorganizing, for example, their writing into a more developed state or, in the parlance of complex dynamic systems theory, when a control parameter causes a new attractor state to emerge (see Larsen-Freeman & Cameron, 2008 for a detailed account of this process), or a combination of these features, that developmental trajectories across learners varied, that writing development was informed by a complex series of language learning affordances, that writing development moved in spurts of progress and moments of regress, and that change occurred at varying rates across learners. In sum, these studies have highlighted that L2 writing development is an amalgam of complex, situated affordances that result in complex, dynamic, and varied instances of development consistent with complex dynamic systems. However, absent in previous studies are analyses of authorial voice construction, a perspective I will later address in Chapter 6 of this thesis.

## 2.5 Summary

This chapter has addressed three key issues relevant to this thesis study. First, pedagogical stylistics research has advocated the potential merit of stylistics for mediating L2 academic writing. Recent studies have attempted to capture its effectiveness in this role, and with a degree of success, but at present there remains a lack of empirical data and corresponding firm evidence to validate its effectiveness in L2 academic writing classrooms. This thesis study addresses this gap in the literature by examining the effectiveness of

pedagogical stylistics in mediating authorial voice development and overall writing development in Japanese learners of English academic writing. Second, while recent studies have highlighted the impact of socio-cultural forces on authorial voice development, including learners' personal histories and how learners acculturate into new discourse spaces, research that traces how learners construct voice at the microgenetic level is limited. This thesis study addresses this shortcoming by exploring the microgenetic processes informing voice development in the essay writing of Japanese ESL writers. Herein, the aim is to reveal cognitive processes informing voice construction and to help consolidate recent findings that can then contribute to better understanding voice development in L2 academic writing contexts. Third, certain valuable insights have widened educators' understanding of concept-development in the tradition of Vygotsky and Galperin. Specifically, CBI has been shown to be useful for advancing language proficiency across different L2 academic writing contexts. In this thesis study I have aimed to expand the basis for advocating CBI's usefulness, and in doing so aimed to further develop the understanding of CBI and concept development in one L2 writing context. Last, this chapter reviewed alternative research approaches to this thesis project and highlighted the relevance of complex dynamic systems theory for interpreting the findings.

## Chapter 3

### Methods

The previous chapter highlighted the need for empirically-based research on pedagogical stylistics and observed the importance of further developing microgenetic understandings of voice construction in L2 academic writing. This thesis study was designed to address these needs. I employed a sequential mixed-method research design that draws on quantitative and qualitative paradigms in “supplementary or complementary forms” (Strauss & Corbin, 1998, p. 34). Drawing from a quantitative-qualitative typology developed in part by Johnson and Christensen (2004) and summarized in Dörnyei (2007), this study interprets the traditional quantitative-qualitative paradigm as follows: QUAN → qual. (The uppercase letters denote the more dominant research focus and the arrow denotes the sequence of the methods.) To complement my primary focus on statistical analyses of learners’ essay writing behaviour, this mixed-method design allowed me to both track emerging themes motivating the choices involved in participants’ writing and to track microgenetic processes informing authorial voice development.

Initially, I attempted to employ a time-series design to collect quantitative data. Accordingly, the classroom tasks were tightly controlled and spaced at equivalent intervals across the intervention period. However, I realized that, even with such controls and multiple, comparable writing tasks, there would be insufficient data points with which to examine the data rigorously in the strict manner of a time-series research design. Nonetheless, the design of the classroom tasks and the data collection points still enabled an examination of how learner development interacts with the independent variable (pedagogical stylistics). Furthermore, rather than focusing on a comparison across different

groups, this design enabled an in-depth analysis of learner writing performance across baseline, treatment, and post-intervention data, including posttests and delayed posttests. (Research question five is an exception to this, as this question involves a comparison between the intervention group [ $n = 7$ ] and the control group [ $n = 3$ ].) At the end of the intervention phase of this study I employed a combined stimulated recall and semi-structured interview to collect verbal data on learners' perspectives of the intervention phase of this study. Alongside statistical data, these perspectives served to reveal the microgenetic processes informing voice development (operationalized as the use of hedges, boosters, attitude markers, authorial self-mention, and direct reader references). This chapter outlines the data collection and analysis phases of my thesis study.

### 3.1 Recruitment

The intervention phase of this study took place in a large Canadian city, and was situated within a three-week TOEFL test preparation writing course offered free to volunteer participants recruited from the local population. Ten adult, L1 Japanese, English students were recruited for this study (my intention was to recruit a larger group of participants, but I was unable to do so). These participants were recruited using flyers (Appendix A) distributed and placed at strategic locations around the city. A free advertisement depicting the flyer was also placed in the announcement section of two websites belonging to online magazines intended for the local Japanese community. As I noted in Section 1.3, my experiences with Japanese learners informed my decision to recruit from this population.

Inclusion criteria for this study were as follows. Participants must: (a) be native speakers of Japanese; (b) be 18 to 26 years old (i.e., preparing for or registered in

undergraduate university programs in English); (c) have a vested interest in studying for the TOEFL exam or a similar standardized English proficiency exam, like the IELTS (either to continue their studies or to further their employment opportunities); (d) have limited experience with stylistics; (e) have intermediate or advanced proficiency in English, approximately equivalent to a score of 68 on the TOEFL iBT or a score of 6 on the IELTS academic, or better; and (f) have previously taken a standardized English proficiency exam that includes an argumentative essay task (e.g., TOEFL, IELTS). This last criterion was included to ensure that participants had experience with the writing genre under consideration—thus helping participants to connect the intervention activities with the goal of improving voice in argumentative essay writing. Regarding the inclusion criteria, three volunteer participants were excluded because their TOEFL iBT scores were considerably lower (54, 56, and 57) than the standard I had set for this study. Substantial experience with stylistic analyses, whether from participants' L1 or elsewhere, was the sole exclusion criterion.

Participants were introduced to the study during a one-hour information session. Participation in this session was mandatory, and consent from interested participants was obtained at the end of the information session. Of the ten participants recruited for this study, seven were randomly selected to participate in the intervention phase of this study. (I employed Microsoft Excel to generate a random table of numbers for this task. This function of Microsoft Excel was used throughout this study in instances where random selection was required.) The remaining three participants formed the control group.

Three outside raters were also recruited for this study. These raters were recruited by a letter (Appendix B) distributed via email over various L2 teaching listservs in the local

community. The inclusion criteria for the outside raters were that the raters: (a) must be trained English language professionals (e.g., possessed a Teachers of English to Speakers of Other Languages [TESOL] Certificate, a Certificate in English Language Teaching to Adults [CELTA] Certificate, or a comparable certificate), or have an advanced degree in the field of second language education (e.g., an MA in Applied Linguistics or TESOL, or a comparable degree); and (b) must have at least three years of experience scoring essays associated with standardized English proficiency tests, such as the TOEFL or the IELTS. I also provided rater training to facilitate inter-rater reliability, which consisted of two sessions totalling four hours, and the training employed widely available TOEFL iBT independent sample essays, across proficiency levels. Alongside me, one of these raters was also employed to code data related to concept development (described in Subsection 3.4.2).

### **3.2 Participant biodata**

Prior to commencing the intervention participants completed an online biodata questionnaire (see Appendix C) administered using eSurveysPro (a free, online questionnaire program). These data helped inform my study by, for example, ensuring participants met the inclusion criteria for this study and by providing relevant background information. These data also provided insight into motivating factors for each participant and a relatively complete profile of each participant. Further, the questionnaire also enquired about English language studies that participants may have been undertaking concurrently with the research that could have influenced the effects of the intervention. This measure, coupled with the sole exclusion criterion (experience with stylistic analyses), was taken to increase the internal validity of the study. These data were also collected to facilitate replication and to help

readers of this study understand the context of my research, as well as to further determine the applicability of the findings to readers in other educational contexts (Mackey & Gass, 2005).

### 3.3 Instruments

During the TOEFL writing course intervention three dependent variables were measured: academic writing proficiency, indicators of voice in academic writing, and the concept development of hedges, boosters, attitude markers, authorial self-mention, and direct reader references. These variables were measured using the following research instruments: the TOEFL iBT independent writing task rubric (Appendix D) (employed to measure academic writing proficiency); a voice rubric adapted from Zhao (2013) (Appendix E) (employed to score the quality of voice in academic writing); a rhetorical feature use frequency tally sheet (Appendix F) (used to track the instances that each rhetorical feature was employed across the writing tasks); worksheets (Appendix G) on stylistic analyses of items from the literary canon (adapted from Lin, 2010), audio recordings of participants discussing their stylistic analyses of items from the literary canon (see Subsection 3.4.2), and essays collected from the TOEFL iBT independent writing task (used to analyze the concept development of the five voice features). Last, a combined stimulated recall and semi-structured interview (Appendix H) was employed to enhance my understanding of learners' perspectives of their experiences regarding the intervention phase of this study as well as to provide insights into the cognitive processes informing their writing. The independent variable was classroom teaching and learning via pedagogical stylistics. A detailed account of how pedagogical stylistics was employed in the classroom is provided in Chapter 4.



Data for scoring writing proficiency, the quality of voice in writing, and the concept development of the five rhetorical features were collected, in part, from responses to essay writing prompts based on the TOEFL iBT independent writing task. This writing task was selected as a data collection tool for three reasons: First, the argumentative writing genre of the TOEFL iBT independent writing task has frequently been identified as a typical writing genre in the university context (Helms-Park & Stapleton, 2003; Reid, 2001; Zhao, 2010, 2013). This writing task is appropriate, then, for measuring writing performance amongst university-aged, international learners familiar with and interested in entering English-dominant university contexts. Second, the argumentative writing genre is also a typical task associated with high-stakes writing (Jeffrey, 2009) such as the TOEFL test. The participants all indicated an interest in performing well in high-stakes tests, so examining their writing performance and development within this context seemed appropriate and also afforded participants the opportunity to benefit from the intervention. Third, previous research (Zhao 2010, 2013) had suggested that the quality of voice manifest in response to the TOEFL iBT independent writing task functions as an indicator of essay scores.

In the opening paragraph of this section, I noted that the voice rubric employed in this study was adapted from Zhao (2013). By employing a confirmatory factor analysis, a principal component analysis, and correlational analyses on 480, double-rated argumentative essays provided by Educational Testing Service (ETS), Zhao developed and validated a voice rubric that measures the strength of authorial voice in, specifically, high-stakes writing contexts such as the TOEFL iBT independent writing task. Given that my thesis study also examines this writing task, I adapted Zhao's voice rubric to measure the quality of voice in the participant writing samples that I collected for this study.

Nine separate sittings of the TOEFL iBT independent writing task formed the dataset examined in this study. The first three written compositions established a baseline, collected in the week prior to beginning the intervention. Four written compositions were also collected during the intervention, on Days 4, 6, 8, and 10. The remaining two samples were collected from a posttest (on Day 12, the final day of the intervention) and a delayed posttest (four weeks after the posttest).

Each writing task was completed in the same computer lab at a local university. The essay tasks featured writing prompts designed by ETS (Educational Testing Service, n.d.), presumed from pretesting and other analyses to be equivalent to each other in content and rhetorical format. Educational Testing Service provides, online and free of charge, approximately 180 sample writing prompts for the independent writing task of the TOEFL iBT. Of these 180 samples, 100 samples were selected as possible options for use in this study. Each of these 100 sample questions asks test takers to take an explicit for-or-against stance, and responses to these types of prompts typically adhere to the genre-model of an argumentative essay (Helms-Park & Stapleton, 2003; Jeffrey, 2009; Reid, 2001; Zhao, 2010). The prompts may ask test takers to agree or disagree with a position, or to defend their preference for one of two positions. For example, “Do you agree or disagree with the following statement? Parents are the best teachers. Use specific reasons and examples to support your answer” (Educational Testing Service, n.d.) or “Some people prefer to get up early in the morning and start the day’s work. Others prefer to get up later in the day and work until late at night. Which do you prefer? Use specific reasons and examples to support your choice” (Educational Testing Service, n.d.). (Other essay prompts allow for a more open-ended response. For example, “If you could meet a famous entertainer or athlete, who

would that be, and why? Use specific reasons and examples to support your choice”

Educational Testing Service, n.d.). The writing prompts employed in this study (Appendix I) were randomly selected from the aforementioned 100 samples.

To assess the concept development of the five rhetorical features (hedges, boosters, attitude markers, authorial self-mention, and direct reader references), two concept development rubrics were employed. These rubrics coded concept development across writing and speaking tasks. Both rubrics coded data from in-class activities wherein participants examined and responded to a piece of modern literature (Table 4.2 in Chapter 4 highlights the literary works examined during the intervention stage of this study). These analyses were performed on stylistic worksheets (Appendix G) that were assigned daily, beginning on Day 3 of the intervention—and barring Day 7 (based on informal student feedback, I concluded on Day 6 that the participants would have appreciated a break from analyzing literary texts). The worksheets were the same throughout the intervention, which strengthened interval validity (Glass et al., 2008).

The stylistic worksheets required that participants perform stylistic analyses of, and engage in discussions on, items that I had selected from the literary canon. The worksheet was divided into three sections: (I) Detect and Identify, (II) Describe, and (III) Discuss. Sections I and II of the worksheet provided data on concept-development associated with identifying the five rhetorical features and explaining their function within the text. Together with responses to the TOEFL essay prompts, Section I and Section II of the worksheet provided data on concept development pertaining to writing. Discussions based on Section III of the worksheet provided data on the relationship between the use of each rhetorical feature of voice in speaking tasks, and the concept development of these features. (I assumed

that capturing data on speaking tasks was relevant because the label of a scientific concept can only be affixed to a concept that has been mastered across different communicative modes.)

### **3.4 Data collection and coding**

Data collection spanned eight weeks. During the first week of the study I collected baseline data. The baseline data consisted of three essays from each participant in response to TOEFL iBT independent writing task essay prompts (the control group only participated in the first of these three baseline writing tasks). Weeks two, three, and four consisted of a classroom intervention. At the end of week four I collected verbal data from interviews with each of the intervention-phase participants. Weeks five, six, and seven did not require learners to participate. At the end of week eight I collected one essay writing task from all ten participants to assess whether there was a lasting effect of the intervention for the participant group, and also to assess any differences between the control group and the participant group.

Data on cognitive uptake regarding voice development was collected from across the dataset. With a few exceptions—for example, statistically significant findings in response to RQ1 could highlight the breadth of cognitive uptake, while findings in response to RQ2 could help determine the path or sequence of cognitive processes—it was impossible to predict which research questions or which element(s) of the dataset would be particularly useful for collecting data on the different dimensions informing the cognitive uptake of voice. Moreover, as I observed in Subsection 2.2.2, there have been only two studies that employed microgenetic analyses to investigate voice development, and so I was unable to

draw from a substantive prior model for collecting data in this context. Accordingly, to increase the validity of this study and to avoid the risk of overlooking data, I attempted where possible to gather extensive data on cognitive change across the entire dataset described below.

### 3.4.1 Collecting and coding data from the TOEFL iBT independent task essay writing samples

I supervised the writing tasks, and participants responded to the writing prompts, using Microsoft Word 2010 on Microsoft Windows-based desktop computers. Prior to the first writing task, and to maximize ecological validity, participants received a short lesson explaining the cut and paste functions on their computers. This lesson ensured that the participants were comfortable with cutting and pasting, as these functions are available to test-takers during the official ETS-sanctioned TOEFL iBT. The spelling and grammar check features were disabled (as in the TOEFL iBT setting), and a clock counting down from 30 minutes—indicating the time remaining to respond to the essay prompt—was displayed on a large screen at the front of the computer lab. Nine essay writing tasks were collected from each of the seven intervention participants (resulting in 63 essays). Two essay writing tasks (the first writing task informing the baseline data and the delayed posttest) were collected from each of the three control-group participants (six essays). In total, 69 essays were collected as this part of the dataset.

After labeling each essay with a random identification code (to conceal the identities of individual participants), the essays were scored by three raters using two rubrics; the raters were also blinded as to the times at which the compositions were written. First, the essays

were scored for the overall quality of academic writing proficiency using the TOEFL iBT independent writing task rubric (Appendix D). Second, a score was also attributed to the essay's overall quality of voice. Voice was scored using a voice rubric adapted from Zhao (2013) (Appendix E). Two coders (I, and one of the other raters employed for scoring the essays) also coded the essays for the frequency of use of each target rhetorical feature (hedge, booster, attitude marker, authorial self-mention, and direct reader reference). To facilitate accurate coding, the essays were reviewed three times by each coder and incidents of each rhetorical feature were recorded on a tally sheet (Appendix F).

Interrater reliability regarding the essay scores and the voice scores was assessed using Fleiss' kappa coefficient (suitable for measuring reliability when there are more than two raters, Fleiss, 1971). Reliability was interpreted using the following scale: 0 to 0.40, poor to fair; 0.41 to 0.60, moderate; 0.61 to 0.80, substantial; and 0.81 to 1.00, almost perfect (following Landis & Koch, 1977). Regarding the quality of writing scores on the TOEFL iBT independent writing task, interrater reliability was  $\kappa = 0.81$ . Regarding the voice scores, interrater reliability was  $\kappa = 0.89$ . Both of these levels of agreement represent an acceptable level of interrater reliability (Landis & Koch, 1977).

Analyses of coder reliability for the frequency of each rhetorical feature in participants' compositions employed Cohen's kappa, as there were only two coders. Coder reliability here was initially  $\kappa = 0.92$ . Where I and the other rater disagreed on how to code an item we returned to discuss the problem until we reached a solution: for example, interpreting some pronouns (e.g., *we*, *us*, and *our*) caused discrepancies in the coding. To solve this issue we decided that pronouns that referred to the author and the reader would be tagged as a direct reader reference. In contrast, pronouns that referred to the author and

another subject in the essay were coded as authorial self-mention. An example of a coding discrepancy we encountered in this vein is from a participant's written composition: "The time that spend with *family* and friend may be more important than work. My father works everyday for *us* [emphases mine] [*sic*]" (Pi, TOEFL1). The referent of "us" is the author and her family, and so we coded the pronoun "us" as an authorial self-mention. After we reviewed the discrepancies in the coding and we worked through each problem, coder reliability was  $\kappa = 1$ .

### 3.4.2 Collecting and coding data regarding concept development

Data concerning the concept development of each rhetorical feature were coded across two modes: writing and speaking. For concept development in the written mode, data were from the essay writing samples and from the stylistic worksheets (Appendix G). For concept development in the spoken mode, data were from the stylistic worksheets. The worksheets were collected daily (barring Day 7), beginning on Day 3 of the intervention. In total, I collected 63 worksheets (seven participants  $\times$  nine worksheets each). From these 63 worksheets, data sampling was employed across an interval scale of every second day (on Days 4, 6, 8, 10, and 12 of the intervention phase of this study). This sampling technique resulted in coding 35 worksheets for analysis (seven participants  $\times$  one work sheet for each of the five days noted above). I first describe how I coded concept development data related to writing.

As noted in Section 3.3, the stylistic worksheet was divided into 3 sections: (I) Detect and Identify, (II) Describe, and (III) Discuss. I employed a framework adapted from Brooks et al. (2010) (Appendix J) to code Section I, Section II, and the essay samples. This

framework was adapted to approximate the development of each rhetorical feature along the concept development continuum (see Section 1.2.3, Table 1.1). For example, if a learning incident was coded as a five, this would indicate that the participant demonstrated scientific concept development of the rhetorical feature in question, whereas a learning incident coded as one would indicate that the concept development of a given rhetorical feature was (approximately) in the early stages of scientific concept development, viz., the stage of a spontaneous concept. Together with the essay samples, Section I and Section II provided data from identification tasks and writing tasks that highlighted concept-development associated with each rhetorical feature. I now discuss how concept development data related to speaking was collected and coded.

Discussions based on Section III of the stylistic worksheet informed data related to speaking and concept development. Discussions were recorded using a digital voice recorder. The recordings were saved in an mp3 format and were collected on days when students performed stylistic analyses (Days 3 to 12, excluding Day 7). I collected recordings of eighteen discussions (two sets of group discussions per class on days when stylistic analyses were performed). Each recording was of the first ten minutes of each discussion; this uniformity strengthened internal validity. In total, I collected three hours of group discussions.

The discussions were coded using a framework for assessing concept development adapted from Brooks et al. (2010) (Appendix K). Similar to how data were coded regarding concept development in the written mode, the coding here was adapted to approximate the development of each rhetorical feature along the concept development continuum (shown in Table 1.1).



Two coders (I and one of the three raters recruited for scoring the TOEFL essay samples) scored the data related to concept development. The essays, the stylistic worksheets, and the audio recordings of the discussions were triple-coded by each coder. Employing Cohen's kappa, coder reliability regarding the dataset related to concept development was initially  $\kappa = 0.92$ . We returned to instances where discrepancies existed with the coding and discussed the problems until we reached a solution. For example, we encountered some ambiguities with our coding that were a result of auditory interference (an interlocutor coughing during someone's utterance). Where we could not confidently interpret the content of an utterance we decided not to code these instances of speech. In total, there were seven instances of speech that we did not code for this reason. After we reviewed all of the discrepancies in the coding, reliability was  $\kappa = 1$  for the concept development dataset.

### 3.4.3 Collecting and coding self-reported verbal data

A combined stimulated recall and semi-structured interview (Appendix H) was also employed to collect verbal data. The interviews were designed to "get at the research participants' perspectives and meaning" (Johnson & Christensen, 2004, as cited by Dörnyei, 2007, p. 173) in order to enhance and inform the other, statistically based findings from this study. These interviews thereby served to triangulate the data by adding another layer to the dataset beyond statistical measures (Mackey & Gass, 2005). The interviews, performed with the seven intervention participants, were conducted three days after the posttest. The interviews were divided into two parts. Part A was a stimulated recall interview in which participants were asked to comment on the rhetorical moves related to voice features that they engaged during the posttest. Part B of the interview was a semi-structured interview.

The aim of Part B was to document participants' views of the intervention period, specifically, the independent variable pedagogical stylistics. Parts A and B also offered the participants opportunities to comment freely on their writing (Part A) or the intervention phase of the study (Part B).

The coding of the verbal data followed an open coding framework (Mackey & Gass, 2005) adopted from Braun and Clarke (2006). The coding process included my becoming familiar with the data (transcribing where necessary, re-reading, and note taking) and creating initial codes (in a systematic manner and across the dataset, and then collating similar codes) (Braun & Clarke, 2006).

### 3.5 Data analyses

In this section I document the data analysis phase of this study. The first subsection is organized by research questions to clearly outline how the data I collected were employed and analyzed. The subsequent subsection outlines how the self-reported verbal data were analyzed.

3.5.1 Research Question 1 (RQ1): *Is there an increase, over the intervention period, in participants' (a) overall quality of writing, and (b) overall quality of voice in writing? Is any increase in the quality of writing and the quality of voice sustained after the intervention?*

After I finished collecting the data, and once the writing samples were coded, I calculated measures of central tendency (means and standard deviations) for the TOEFL iBT writing scores and the voice scores. I calculated these measures for each participant and for

the group. I then employed scatterplots to visually examine the data for patterns. After consulting with a research systems analyst at my home institution regarding testing-methods, a Friedman test was selected to investigate any statistical differences across the written compositions regarding the writing scores and the voice scores. Following Sheskin (2003) the effect size statistic was calculated using Kendall's coefficient of concordance (Sheskin notes that Kendall's coefficient of concordance can be used as a measure of effect size when the Friedman test is used to investigate statistical difference—see Test 25, Section VII of Sheskin). The effect size was interpreted using the same scale as the eta squared statistic (“Comparison of effect size”, n.d.). Accordingly, 0.02 or less was interpreted as a small effect size, 0.13 and related results were interpreted as a medium effect size, and 0.26 or greater was interpreted as a large effect size (Watson, 2013).

Where a statistically significant difference was detected, post hoc analyses with Wilcoxon signed-rank tests were conducted with a Bonferroni correction applied (Field, 2009; Larson-Hall, 2010). The Bonferroni correction was applied despite the risk of a Type II error (Cabin & Mitchell, 2000; Field, 2009; Larson-Hall, 2010; Perneger, 1998). To offset this risk, and following Biddix (n.d.), Cabin and Mitchell (2000), Ellis (2010), and Perneger (1998), I emphasized the effect sizes in reporting the results.

The effect sizes from the post hoc Wilcoxon signed-rank tests were calculated using the following formula:  $r = Z/\sqrt{N}$  (Field, 2009; Rosenthal, 1991). I then converted the resultant  $r$  coefficient into the Cohen's  $d$  index, based on Ellis' (2010) notion that both indices are standardized and “can be transformed into the other” (Ellis, 2010, p. 16; see also Ellis, 2010; Furr, 2008; Rosenthal, 1994; Salekin, Rogers & Sewell, 1996). The  $r$  coefficient was converted to report a standardized effect size (Dörnyei, 2007; Ellis, 2010) to evaluate

“results across samples, designs and analyses (Dörnyei, 2007, p. 212). Given that Cohen’s  $d$  is a standard measure of effect size (Dörnyei, 2007; Mackey & Gass, 2005) and that Cohen’s  $d$  more frequently appears as a measure of effect size in applied linguistics research, standardizing the  $r$  coefficient into the Cohen’s  $d$  index seemed appropriate in the context of this study.

Following Cohen (1988) and Rosenthal (1994), this conversion used the Rosenthal transformation:  $d = 2r \div \sqrt{1 - r^2}$ . However, there is a small chance that such conversions minimally underestimate the effect size (Salekin, Rogers, & Sewell, 1996). To control for a Type II error, then, I reported  $d$  as a matter of convenience for cross-study comparisons, and interpret the effect sizes in this study according to the  $r$  coefficient. This procedure was repeated for RQ1 and RQ5, where the Rosenthal transformation was also employed.

The effect size based on the  $r$  coefficient was interpreted as follows: .29 is a small effect size; .3 to .59 is a medium effect size; .6 to .79 is a large effect size, and .8 or greater is a very large effect size. This scale is an amalgam of interpretive scales designed for the  $r$  coefficient and the Spearman’s rho ( $r_s$ ) coefficient that I later employ in analyses for RQ2, RQ3, and RQ4. As a matter of convenience, these two coefficients have been interpreted using this same scale. This scale also closely resembles the metric Plonsky and Oswald (2014) recently proposed for interpreting effect sizes in L2 research: .25 as small, .40 as medium, and .60 as large.

The statistical analyses thus described were performed to detect any significant increase, and the corresponding effect size, in the overall writing scores and the quality of voice scores across time as a result of the independent variable, pedagogical stylistics.

To determine whether any increase in the quality of writing and the quality of voice was sustained over time (the second part of RQ1), I conducted a Wilcoxon signed-rank test. This test analyzed the data from the posttest and the delayed posttest. I calculated the effect size by using the same formula noted above:  $r = Z/\sqrt{N}$  (Field, 2009; Rosenthal, 1991). As with RQ1, the resultant  $r$  coefficient was also converted to the Cohen's  $d$  index using the Rosenthal transformation.

3.5.2 Research Question 2 (RQ2): *What is the relationship between the frequency of use of each target rhetorical feature of voice, and (a) the overall quality of writing and (b) the overall quality of voice?*

To analyze the relationship between the frequency of use of each target rhetorical feature of voice and the overall quality of writing and the overall quality of voice I calculated Spearman's rank correlation coefficients (i.e.,  $r_s$ ). Prior to calculating Spearman's rho, I calculated measures of central tendency (means and standard deviations) for the frequency tally for each rhetorical feature. Correlations were then examined in two separate instances. First, I examined the correlation of the group mean usage of each rhetorical feature with the group mean essay scores from the TOEFL iBT independent writing tasks. Second, I examined the correlation of the group mean usage of each rhetorical feature with the group mean voice scores collected from the TOEFL iBT independent writing tasks. Following Field (2009) and Slate and Rojas-LeBouef (2011), I interpreted the resultant coefficient as representative of the effect size. I also interpreted the effect size in the same manner as RQ1.

To create alternative representations of the resultant relationships, I employed two different graph types. First, scatterplots were employed to examine patterns of rhetorical

feature use across the written compositions, based on group means. Second, radar graphs were employed to depict individual tendencies in this context. One radar graph was created for each participant in the intervention group. The radar graphs visually document how each individual altered their use of the rhetorical features of voice across all of the writing samples.

### 3.5.3 Research Question 3 (RQ3): *What is the relationship between (a) the quality of writing and (b) the quality of voice in writing?*

To examine the relationship between the quality of writing and the quality of voice in writing I analyzed data that emerged from the two rubrics employed to score each of these variables. The relationship between these two variables was calculated using Spearman's rho. The resultant coefficient was also treated as an indicator of the effect size, following Field (2009) and Slate and Rojas-LeBouef (2011). I interpreted the effect sizes using the same metric that I employed in RQ1.

### 3.5.4 Research Question 4 (RQ4): *What is the relationship between (a) the frequency of use of each target rhetorical feature of voice and (b) the concept development of each rhetorical feature of voice?*

To investigate the relationship between the frequency of use of each target rhetorical feature of voice and the concept development of each rhetorical feature of voice I calculated Spearman's rho by employing two sets of data: the data collected from the rhetorical feature use frequency tally sheet, and the data collected from the concept development scores attributed to the concept development of each rhetorical feature across

writing and speaking tasks. This calculation produced ten coefficients: five coefficients represent the relationship between the use of each rhetorical feature and the concept development of writing; the other five coefficients represent the relationship between the use of each rhetorical feature and the concept development of speaking. I interpreted the resultant coefficients as a marker of the effect size (Field, 2009; Slate & Rojas-LeBouef, 2011). I also interpreted the effect size using the same metric that was noted in RQ1.

3.5.5 Research Question 5 (RQ5): *Are there differences in (a) the quality of writing and (b) the quality of voice in writing between the control group and the course participants?*

As noted in Subsection 3.4.1, the control group only performed two tasks: on the first day of the study when baseline data was collected, and alongside the course participants, the control group responded to the same TOEFL essay prompt as the course participants; on the last day of the study—eight weeks later and again alongside the course participants—the control group responded to the same TOEFL essay prompt as the course participants. The data collected from the control group were subjected to the same analysis as the course participants regarding the quality of writing and the quality of voice in writing. To answer RQ5, I compared the quality of writing and the quality of voice in writing across both groups.

To examine whether there was a significant difference between the control group and the course participants I performed a Mann-Whitney U test. I ran analyses using the first written composition that both groups composed and the last written composition collected at the delayed posttest. I compared the participants against the control group at two points

regarding both essay and voice scores. This comparison resulted in four different Mann-Whitney U tests, wherein I compared the participants against the control group:

- (a) with the initial writing sample, examining the overall quality of the writing;
- (b) with the delayed posttest writing sample, examining the overall quality of the writing;
- (c) with the initial writing sample, examining the overall quality of voice; and
- (d) with the delayed posttest writing sample, examining the overall quality of voice.

The means and standard deviations were also calculated for the overall quality of writing score and the overall quality of voice score. I calculated the effect size as follows:  $r = z/\sqrt{N}$  (Field, 2009, Rosenthal, 1991). I also converted the resultant  $r$  coefficient into the Cohen's  $d$  index in the same manner as RQ1, using the Rosenthal transformation ( $d = 2r \div \sqrt{1 - r^2}$ ). As with RQ1, I converted the  $r$  coefficient to report a standardized effect size; the effect size was interpreted using the same scale described for RQ1.

### 3.5.6 Self-reported verbal data

As noted in Subsection 3.4.3, the verbal data collected from the seven intervention participants were analyzed using an open coding framework (Mackey & Gass, 2005) adopted from Braun and Clarke (2006). Once the data were coded, a thematic analysis (Pavlenko, 2007) of the verbal data ensued. Thematic was operationalized in this study following Braun and Clarke (2006), who defined a theme as important data “in relation to the research question, and [which] represents some level of patterned response or meaning within the data set” (p. 82). (A similar description can be found in Mackey & Gass, 2005.) The thematic analysis of the verbal data engaged the following steps:



- (a) Probing for themes (collecting and organizing data into potential themes):
- (b) Revising themes (reviewing that themes match the initial codes);
- (c) Naming the themes (refining and reviewing the themes against their description); and
- (d) Selecting appropriate data to report (ensuring that the data reported corresponds to the research questions) (Braun & Clarke, 2006).

### 3.6 Summary

This chapter documented the methods employed in the data collection, coding, and analysis phases of this study. Relative to each research question, Table 3.1 provides an overview of the instruments that were employed in this study as well as the data that was collected and how these data were analyzed.

In summary, the total data informing this thesis study included 69 essay responses to TOEFL iBT independent writing task essay prompts (63 from the seven classroom participants and six from the three control participants), 63 stylistic worksheets (nine from each intervention-group participant), three hours of recorded classroom-based participant discussions, and approximately seven hours of interview data (approximately one hour from each intervention participant).

**Table 3.1**  
*Summary of Research Questions, Research Instruments, and Data Analyses\**

Research question	Instruments and data to be analyzed	Analyses
1. Is there an increase, over the intervention period, in participants' (a) overall quality of writing, and (b) overall quality of voice in writing? Is any increase in the quality of writing and the quality of voice sustained after the intervention?	(a) scores generated from the TOEFL iBT independent writing task rubric (Appendix D) (b) scores generated from the voice rubric (Appendix E)	Friedman test (with a post hoc Wilcoxon signed-rank test when necessary) 1a. Wilcoxon signed-rank test
2. What is the relationship between the frequency of use of each target rhetorical feature of voice, and (a) the overall quality of writing and (b) the overall quality of voice?	- tallies generated from the framework for measuring the frequency of use of the rhetorical features of voice in the writing task (Appendix F) (a) scores generated from the TOEFL iBT independent writing task rubric (Appendix D) (b) scores generated from the voice rubric (Appendix E)	Spearman's rank correlation coefficient Radar graphs
3. What is the relationship between (a) the overall quality of writing and (b) the overall quality of voice in writing?	(a) scores generated from the TOEFL iBT independent writing task rubric (Appendix D) (b) scores generated from the voice rubric (Appendix E)	Spearman's rank correlation coefficient
4. What is the relationship between (a) the frequency of use of each target rhetorical feature of voice and (b) the concept development of each rhetorical feature of voice?	(a) tallies generated from the framework for measuring the frequency of use of the rhetorical features of voice in the writing task (Appendix F) (b) scores generated from the coding framework for concept development of rhetorical features of voice in writing (Appendix J) based on the stylistic worksheet and the TOEFL writing task, and then the coding framework for concept development of rhetorical features of voice in speaking tasks (Appendix K) derived from in-class group discussions based on the stylistic worksheet (Appendix G)	Spearman's rank correlation coefficient Line graphs
5. Are there differences in (a) the quality of writing and (b) the quality of voice in writing between the control group and the course participants?	(a) scores generated from the TOEFL iBT independent writing task rubric (Appendix D) (b) scores generated from the voice rubric (Appendix F) (the scores for (a) and (b) are taken from the first writing task and the delayed post-test writing task)	Mann-Whitney U test

\* The verbal data served a supplementary role that informed all of the research questions. Accordingly, given the wide scope of the interview data, the role of the interviews is not included in Table 3.1. Similarly, in the present study it was impossible to predict which elements of the dataset would provide data regarding microgenetic changes in learner cognitive behaviour; accordingly, the exact research questions and instruments employed to track microgenetic changes are not included in Table 3.1.

## Chapter 4

### Instructional Context

This classroom-based study employed a computer lab and a classroom at a large Canadian university for collecting data and for classroom activities. Congruent with principles of CBI, a classroom context is suited for “generating a type of knowledge that is abstract, generalizable, meaningful, systematic, and functional, that is, theoretical, scientific concepts” (Yanez-Prieto, 2010, p. 66). In addition, the context for this classroom-based study was based on ubiquitous TOEFL test preparation courses offered in Canada and Japan. A Google search, for example, conducted in March 2013 with the search term “TOEFL test preparation classes Toronto” revealed 10 courses available in the local community during the spring of 2013, based only on the first page of the search results. The same search using Tokyo in place of Toronto conducted in English and Japanese, revealed similar results for courses in Tokyo for the same time period. Accordingly, this classroom-based design seemed appropriate for both operationalizing CBI and for ecological validity.

To offer a detailed description of the instructional context for this study, in the remaining subsections of this chapter I outline my role as the researcher and instructor in this context and outline the learners’ profiles in detail. I then conclude by documenting the content and sequence of the intervention phase of this study and by providing a narrative account of the teaching-learning interactions.

#### 4.1 Researcher-cum-instructor

Due to issues that surfaced during the pilot phase of this study (e.g., questions that caused the class to veer off topic and thus exposed time constraints), concerns about internal validity became greater than concerns about researcher bias. Furthermore, given the academic background required to effectively teach stylistics using CBI, and the time commitment involved in the study, finding an instructor competent and committed to this project would have been arduous, if not impossible. Accordingly, to control for variables that could contaminate the data, I decided to position myself as researcher-cum-instructor. To account for the instructional context of this study, then, I begin by outlining my experiences as an instructor teaching academic writing to English language learners (ELL) and my experiences employing stylistics as a teaching tool with different L1 Japanese ELLs.

My experiences teaching English academic writing and employing stylistics in a Japanese university (2008-2011) informed my role as the instructor for this study, as suggested already in Section 1.3. First, I have repeatedly taught the argumentative essay genre—the target writing task for this study—to Japanese learners in Japan and Canada. This experience helped me better understand the academic writing needs of Japanese university students. Second, my experiences from 2008 to 2011 helped me refine how I present and operationalize stylistics in the classroom. For example, initial attempts at stylistic analyses of short stories proved overly demanding for students. Students had little interest in exploring texts that appeared removed from their experiences. To offset lack of interest, I now introduce in my classes stylistic analyses with song lyrics selected by students (a tool for stylistic analyses addressed by Wales, 2006). For example, to introduce song lyrics, students are given a copy of the lyrics and the song is played for them twice. Students then proceed to

work through their stylistic analyses using the same process of analysis and the same stylistic worksheet (Appendix G) used for analyzing poems or short stories as outlined in Sections 4.4, 4.5, and 4.7. Introducing stylistics in this manner has greatly enhanced students' motivation and uptake. With stylistic analyses of song lyrics as a base, my experiences also demonstrated that students are then more willing to analyze traditional items from the English literary canon, like poetry or short stories. Together, these experiences teaching academic writing and employing stylistics prepared me for and aided me in the role of instructor during this study.

## **4.2 Biodata information**

The ten randomly selected participants in this study completed a biodata questionnaire (Appendix C). Table 4.1 provides details for each participant (using pseudonyms to protect their privacy).

Of the participants selected for this study Tom, Am, and Pi did not meet all of the inclusion criteria for this study. Tom was 27 years old, and one year older than the age criterion set for this study. The age criterion was designed to attract university-aged students or students who had recently graduated from university. Tom began university one year later than is typical in Japan, and although beyond the age criterion by one year, she had recently graduated from university. Accordingly, I deemed that her participation would not adversely affect my study. Regarding Am and Pi, these participants were included despite not meeting the minimum English language proficiency level outlined in Section 3.1. However, based on the standard error of measurement for the TOEFL iBT (10.56) (Educational Testing Service, 2005) I deemed their TOEFL iBT scores of 62 to be approximately equivalent to the

minimum criterion set of 68. Accordingly, I treated the participants as a homogenous group concerning their language proficiency (in Subsection 6.6.1 I discuss some of the outcomes of this decision with a focus on Am).

The participants all expressed a desire to learn English as the primary reason for being in Canada. All of the participants were also either university or college students in Japan who had taken a break from their university studies or who had just recently graduated. Five of the participants were engaged in language learning with the hopes of securing employment that required higher levels of English proficiency than they currently possessed. Three participants had plans to attend graduate school (two in the United States and one in Canada). Of the ten participants, eight were female and two were male. The mean age of the participants was 23.3.

Each participant had previously written the TOEFL iBT or the IELTS academic, either in Japan or Canada. The mean English language proficiency level for the participants, based on self-reported TOEFL iBT and IELTS academic scores, was 71.9. (Those participants who reported IELTS academic scores had their IELTS academic scores converted to TOEFL iBT scores. This conversion was based on Educational Testing Service, 2011.) Only one participant (Angela) was studying English concurrently at a private language school. The influence of this participant's extra-curricular course work is described in Subsection 6.1.4. Table 4.1 provides a detailed profile of each participant.

As noted above, seven of the ten participants were randomly selected to participate in the three-week, classroom-based intervention that informed this study. The remaining three participants formed the control group. The content, format, and rationale for the intervention are described in the next section.

**Table 4.1**  
*Participant Biodata*

Participant pseudonym	Gender	Age	Undergraduate major	Latest TOEFL iBT score <sup>a</sup>	Purpose for participating	Long-term (career) goals
Pi	F	21	Culture & information	62	Improve essay writing in English & Japanese	Secure employment in Japan that requires advanced English proficiency
Anny	F	21	Linguistics	86 (6.5)	Improve essay writing	Graduate school in the United States
Maru	M	22	Accounting	69 (6)	Improve English to travel	–
Anne	F	24	Media communication	68	Improve overall academic English proficiency	Secure employment in, and immigrate to, Canada
Tom	F	27	Education	79	Improve overall academic English proficiency	Graduate school in Canada
Am	F	25	English literature	62	Improve TOEFL score	Secure employment in Japan that requires advanced English proficiency
Angela	F	21	English literature	72	Improve her TOEFL score so that she can qualify to be an exchange student in the United States	Secure employment in Japan that requires advanced English proficiency
Jill <sup>b</sup>	F	24	Business	79	Improve her English to secure employment in Canada	Secure employment in, and immigrate to, Canada
Erika <sup>b</sup>	F	26	International relations	77	Improve TOEFL score	Graduate school in the United States
John <sup>b</sup>	M	22	Business	65	Improve her English to secure employment in Canada	Immigrate to Canada

<sup>a</sup> Participants who reported IELTS academic scores had these scores converted to TOEFL iBT scores based on ETS (2011); the score in parentheses is the original IELTS academic score reported by the participant.

<sup>b</sup> Participants forming the control group.

### 4.3 Instructional phase

This study employed a CBI, classroom-based teaching method as a backdrop for exploring the research questions. I collected baseline data in the week prior to beginning the intervention; the intervention occurred over three successive weeks (Monday to Thursday, with each class period lasting 80-minutes), totaling 12 days (or 16 hours) of classroom-based intervention.

The intervention was based on Galperin's STI (1969, 1992), a mode of operationalizing Vygotsky's theory of spontaneous and scientific concepts. Systemic-theoretical instruction is based on three key principles: (a) concepts are the unit of instruction; (b) materializing the concepts; and (c) languaging—defined as “the process of making meaning and shaping knowledge and experience through language [via speaking or writing (Swain et al., 2009)]” (Swain, 2006, p. 98)—to assist learners in internalizing the target concepts (Lantolf & Thorne, 2006). These principles are divided into six teaching stages:

1. Motivational stage: introducing learners to the action required to achieve the goal;
2. Orienting stage: presenting learners with the orienting basis of action (OBA);
3. Materialized stage: mastering the action using material or materialized objects;
4. Stage of overt speech: mastering the action at the level of overt speech;
5. Stage of covert speech: mastering the action at the level of speaking to oneself; and
6. Mental stage: transferring the action to the mental level. (Johnson, 2008)

Table 4.2 provides a synopsis of the how the content of the three-week intervention afforded participants opportunities to work through each of these six stages. Table 4.2 also explicitly



Table 4.2  
*Instructional Procedures, Aims, and Rationale*

Class day	Instructional procedures	Aim(s)	Theoretical rationale	Classroom activities
Day 1	<p>(a) introduced stage one of Galperin's STI<sup>a</sup></p> <p>(b) introduced stage two of STI (the OBA<sup>b</sup>) which outlines the conceptual units—the five rhetorical features—that are the focus of this study</p>	<p>(a) to provide students with an explicit understanding of the goals of the class and the steps required to achieve the goals</p> <p>(b) to provide students with a visual map to aid their concept development</p>	<p>(a) explicit instruction is central to Galperin's STI for students to develop scientific concepts (herein referred to as 'explicit instruction')</p> <p>(b) a visual map (the OBA) orients students to how the goals (the scientific concept development of the five rhetorical features of voice) can be achieved</p>	<p>(a) in two groups students, and then collectively with the instructor, discussed the goals for the class; the class also discussed the steps required to achieve the goals</p> <p>(b) first in two groups, and then collectively, discussions about the OBA were held regarding the function and utility of the conceptual units presented in the OBA and the OBA's potential usefulness for learning</p>
Day 2	<p>(a) discussed the TOEFL iBT independent writing task scoring rubric</p> <p>(b) reviewed an ETS video describing how to successfully navigate the independent writing task</p> <p>(c) introduced stylistics and stylistic analyses</p> <p>(d) group stylistic analysis of Bob Dylan's song "Oxford Town"</p>	<p>(a) to provide students with an explicit understanding of how their writing is being scored</p> <p>(b) to provide students with a mediating tool that aids concept development</p> <p>(c) to explain and provide students with the main mediating tool for developing the conceptual units</p> <p>(d) to provide students with a model of a stylistic analysis and to make explicit the expectations of the task</p>	<p>(a) explicit instruction</p> <p>(b) the introduction of stage three of Galperin's STI (a mediating tool materializes the conceptual units, thus aiding concept development)</p> <p>(c) explicit instruction</p> <p>(d) providing students opportunities to develop conceptual units through stages four and five of STI improves likelihood of scientific concept development</p>	<p>(a) in two groups students, and then collectively with the instructor, discussed the descriptors for each score; sample questions for discussion included "What does 'coherence' mean in essay writing?"</p> <p>(b) students discussed the content of the video in two groups, followed by a class discussion</p> <p>(c) students listened to a lecture on the value and potential of stylistics for language learning; students worked through examples of texts demonstrating stylistic analyses</p> <p>(d) led by the instructor, the class worked through a model stylistic analysis, using the worksheet (Appendix G) as a guide</p>

Day 3 <sup>c</sup>	<p>(a) examined developing and organizing ideas regarding the TOEFL iBT independent writing task</p> <p>(b) used the OBA to help examine sample essay paragraphs</p> <p>(c) reviewed how to successfully perform a stylistic analysis and reviewed further possible analyses of Bob Dylan's song "Oxford Town"</p>	<p>(a) to provide students with concrete ideas that match the scoring rubric</p> <p>(b) to provide students with continued access to the OBA and to demonstrate its continued relevance to their learning</p> <p>(c) to provide students with opportunities to better understand how to perform stylistic analyses</p>	<p>(a) explicit instruction</p> <p>(b) explicit instruction and providing opportunities to develop conceptual units through stages four and five of STI improves likelihood of scientific concept development</p> <p>(c) explicit instruction and providing opportunities to develop conceptual units through stages four and five of STI improves likelihood of scientific concept development</p>	<p>(a) students examined model essays for development and organization, and then examined their essays that were collected as part of the baseline data; discussions ensued</p> <p>(b) sample paragraphs from essays provided by ETS were distributed; individually and then in pairs, and finally as a group the class discussed how the OBA informs the choices made by the author regarding the five rhetorical features</p> <p>(c) individually students re-examined "Oxford Town"; they performed the analysis without the aid of the previous day's notes; as a class we discussed the analyses, highlighting positive and negative elements of analyses</p>
Day 4 <sup>c d</sup>	<p>(a) stylistic analysis of David Erico's short story "Me and You"</p> <p>(b) reviewed possible stylistic analyses of David Erico's short story "Me and You"</p>	<p>(a) to provide students with their first opportunity to both do their own stylistic analysis and to develop their language awareness regarding the conceptual units</p> <p>(b) to provide students with immediate feedback on their stylistic analysis</p>	<p>(a) providing opportunities to develop conceptual units through stages four and five of STI improves likelihood of scientific concept development</p> <p>(b) providing graduated feedback based on interactionist dynamic assessment (IDA) and Aljaafreh and Lantolf's (1994) regulatory scale aids development of conceptual units</p>	<p>(a) students were given a copy of the short story and the worksheet; they worked through their stylistic analysis individually, and then in two groups they discussed their analyses</p> <p>(b) as a class we discussed different possible stylistic interpretations of the text and discussed comments made during the group discussions</p>
Day 5 <sup>c</sup>	<p>(a) provided feedback on writing task one, specifically regarding the use of the five rhetorical features</p>	<p>(a) to provide students with comments on their essays relevant to the conceptual units to improve their writing</p>	<p>(a) providing graduated feedback based on IDA and Aljaafreh and Lantolf's (1994) regulatory scale aids development of conceptual units</p>	<p>(a) I returned ungraded essays and provided prompts in the margins where corrections could be made; I walked around the class and provided more explicit prompts when required</p>

	(b) provided Visual Aid II (Appendix M) that materializes the conceptual units (another mediating tool)	(b) to provide students with another mediating tool related to stage three that aids concept development	(b) a mediating tool materializes the conceptual units and affords another opportunity to aid concept development	(b) first in two groups, and then as a class, discussions about Visual Aid II were held regarding its utility as a mediating tool for concept development
	(c) discussed typical elements of, and provided practice with, pre essay writing tasks (outlining)	(c) to provide students with a tool to develop their ideas in the pre-writing stage of essay writing	(c) explicit instruction	(c) students were provided with different models of outlining, and were given sample essay questions from ETS on which to practice outlining
	d) stylistic analysis of Chairlift's song "Bruises"	d) to provide students with continued opportunities to develop their language awareness regarding the conceptual units	d) providing continued opportunities to develop conceptual units through stages four and five of STI improves likelihood of scientific concept development	d) students were given a copy of the song lyrics and the song was played for them twice; then they worked through their stylistic analysis individually, and then in two groups they discussed their analyses
Day 6 <sup>c,d</sup>	(a) stylistic analysis of Charles Bukowski's poem "trashcan lives".	(a) to provide students with continued opportunities to develop their language awareness regarding the conceptual units	(a) providing continued opportunities to develop conceptual units through stages four and five of STI improves likelihood of scientific concept development	(a) students were given a copy of the poem and the worksheet; they worked through their stylistic analysis individually, and then in two groups they discussed their analyses
Day 7	(a) provided feedback on writing task two  (b) used the OBA and Visual Aid II to work on exercises highlighting voice in academic writing  (c) examined how to develop strong topic sentences for	(a) to provide students with comments on their essays relevant to the conceptual units to improve their writing  (b) to provide students with continued access to the OBA and Visual Aid II and to demonstrate its continued relevance to their learning  (c) to provide students with a	(a) providing graduated feedback based on IDA and Aljaafreh and Lantolf's (1994) regulatory scale aids development of conceptual units  (b) explicit instruction and providing opportunities to develop conceptual units through stages four and five of STI improves likelihood of scientific concept development	(a) I returned ungraded essays and provided prompts in the margins where corrections could be made; I walked around the class and provided more explicit prompts when required  (b) sample paragraphs from essays provided by ETS were distributed; individually and then in pairs, and finally as a group the class discussed how the OBA informs the choices made by the author regarding the five rhetorical features

	<p>paragraphs based on ETS-published materials</p> <p>(d) reviewed a TOEFL iBT independent writing task checklist</p>	<p>tool to develop their ideas in the pre-writing stage of essay writing</p> <p>(d) to provide students with a tool to develop their ideas in the pre-writing stage of essay writing</p>	<p>(c) explicit instruction</p> <p>(d) explicit instruction</p>	<p>(c) sample paragraphs from essays provided by ETS were distributed; individually and then in pairs, and finally as a group the class discussed how the topic sentences function to support the development of the essay; student were then given sample essay questions from ETS with which to practice topic sentences</p> <p>(d) we reviewed the content of the checklist and discussed the relevance of each item</p>
Day 8 c d	<p>(a) stylistic analysis of Ray van Horn Jr.'s short story "Banished to My Room"</p>	<p>(a) to provide students with continued opportunities to develop their language awareness regarding the conceptual units</p>	<p>(a) providing continued opportunities to develop conceptual units through stages four and five of STI improves likelihood of scientific concept development</p>	<p>(a) students were given a copy of the short story and the worksheet; they worked through their stylistic analysis individually, and then in two groups they discussed their analyses</p>
Day 9 c	<p>(a) provided feedback on writing task three</p> <p>(b) provided participants with exercises to review their own writing from writing task three (exercises were focused on introductions and conclusions and included the OBA and Visual Aid II)</p> <p>(c) stylistic analysis of U2's song "Sunday Bloody Sunday"</p>	<p>(a) to provide students with comments on their essays relevant to the conceptual units to improve their writing</p> <p>(b) to provide students with feedback and continued opportunities to improve their overall academic writing performance, with a particular focus on the conceptual units</p> <p>(c) to provide students with continued opportunities to develop their language</p>	<p>(a) providing graduated feedback based on IDA and Aljaafreh and Lantolf's (1994) regulatory scale aids development of conceptual units</p> <p>(b) explicit instruction</p> <p>(c) providing continued opportunities to develop conceptual units through stages four and five of STI improves likelihood of scientific concept development</p>	<p>(a) I returned ungraded essays and provided prompts in the margins where corrections could be made; I walked around the class and provided more explicit prompts when required</p> <p>(b) students were given the introduction and conclusion they wrote on the previous day; they were also given model introductions and conclusions provided by ETS; exercises were designed to help students notice the differences in their writing and the models, and were asked to rewrite their work to make it stronger; students were also asked to employ the OBA and/or Visual Aid II to help them</p>

		awareness regarding the conceptual units		(c) students were given a copy of the song lyrics and the song was played for them twice; then they worked through their stylistic analysis individually, and then in two groups they discussed their analyses
Day 10 <sup>c d</sup>	- stylistic analysis of Charles Bukowski's poem "drive through hell"	(a) to provide students with continued opportunities to develop their language awareness regarding the conceptual units	(a) providing continued opportunities to develop conceptual units through stages four and five of STI improves likelihood of scientific concept development	(a) students were given a copy of the poem and the worksheet; they worked through their stylistic analysis individually, and then in two groups they discussed their analyses
Day 11 <sup>c</sup>	(a) provided feedback on writing task four  (b) provided participants with an exercise to review how—with the aid of the OBA and Visual Aid II—the conceptual units contribute to good writing  (c) reviewed possible analyses of Bukowski's poem "drive through hell"  (d) stylistic analysis of Margaret Atwood's poem "Doorway"	(a) to provide students with comments on their essays relevant to the conceptual units to improve their writing  (b) to provide students with continued access to the OBA and Visual Aid II and to demonstrate its continued relevance to their learning  (c) to provide students with feedback on their stylistic analysis  (d) to provide students with continued opportunities to develop their language awareness regarding the conceptual units	(a) providing graduated feedback based on IDA and Aljaafreh and Lantolf's (1994) regulatory scale aids development of conceptual units  (b) explicit instruction and providing opportunities to develop conceptual units through stages four and five of STI improves likelihood of scientific concept development  (c) providing graduated feedback based on IDA and Aljaafreh and Lantolf's (1994) regulatory scale aids development of conceptual units  (d) providing continued opportunities to develop conceptual units through stages	(a) I returned ungraded essays and provided prompts in the margins where corrections could be made; I walked around the class and provided more explicit prompts when required  (b) Using an anonymous essay from TOEFL essay task four, students discussed why the essay was good and made suggestions for further improving the essay; the students were asked to ground their suggestions in the OBA and/or Visual Aid II  (c) as a class we discussed different possible stylistic interpretations of the text and discussed comments I collected from the previous day's group discussions  (d) students were given a copy of the poem and the worksheet; they worked through their stylistic analysis individually, and then in two groups they discussed their analyses

			four and five of STI improves likelihood of scientific concept development	
Day 12 <sup>c,d</sup>	(a) stylistic analysis of Margaret Atwood's poem "This is a Photograph of Me"	(a) to provide students with continued opportunities to develop their language awareness regarding the conceptual units	(a) providing continued opportunities to develop conceptual units through stages four and five of STI improves likelihood of scientific concept development	(a) students were given a copy of the poem and the worksheet; they worked through their stylistic analysis individually, and then in two groups they discussed their analyses

<sup>a</sup> Systemic theoretical instruction.

<sup>b</sup> Orienting basis of action.

<sup>c</sup> Days that stylistic analysis worksheets were collected and days when discussions were audio recorded for data collection.

<sup>d</sup> Days wherein students answered essay prompts for data collection, and days where the stylistic worksheets and audio recordings that were collected were subject to analysis as part of a data-sampling regime.

states the aim of each task alongside the theoretical rationale for the task. A brief description of how activities were executed by the participants is also provided.

Below I outline the ways in which stylistic analyses were employed in the context of this study, and I document how stylistics was operationalized in the classroom.

#### **4.4 Teaching with stylistics**

While discussing teaching methods commonly associated with stylistics, Bellard-Thomson (2011) highlighted the “eclectic nature of stylistics” (p. 55) and how this diversity contributes negatively to defining the “content of stylistics courses with any certainty” (p. 55). Given the absence of any one standardized course in stylistics (as no one standard method for applying stylistics exists, Bellard-Thomson, 2011 and Jeffries, 2000), I drew on the narrow focus of this study’s intervention—improving the quality of voice in academic essay writing—to limit the range of stylistic analysis. This research examined only one stylistic or linguistic feature of writing that permeates academic and non-academic texts: voice. More specifically, the stylistic analyses performed in class—following Bex’s (1999) suggestions for engaging students in classroom-based stylistic analyses, and adapted from a framework provided by Verdonk (2013), Short (1996), and Leech and Short (2007) for examining the linguistic structure of a text—asked participants to consider how authors’ lexical choices inform the semantic and pragmatic features of the text associated with the five target rhetorical devices. Accordingly, these five rhetorical devices featured throughout the intervention as the focal point for the stylistic analyses of the literary texts. The texts that the intervention group examined consisted of the lyrics to three songs (Bob Dylan’s “Oxford Town”, 1962, Chairlift’s “Bruises”, 2008, and U2’s “Sunday Bloody Sunday”, 1983), two

short stories (David Erico’s “Me and You”, 2009 and an adaptation of Ray van Horn Jr.’s “Banished to My Room”, 2000), and four poems (Charles Bukowski’s “trashcan lives”, 2002 and “drive through hell”, 2002, and Margaret Atwood’s “Doorway”, 1984 and “This is a Photograph of Me”, 1998). The last column in Table 4.2 provides descriptions of how each of these texts was analyzed. For example, on Day 6 students were introduced to Charles Bukowski via a short 5-minute presentation that I conducted. The presentation provided some information about the poet and the social context in which he wrote. Students were then given a copy of Bukowski’s poem “trashcan lives” and were asked to complete the stylistic worksheet (Appendix G) that concluded with a group discussion (Section 4.7 also provides a more detailed, narrative account of how these texts were presented, analyzed, and discussed in class). During analyses of these texts participants underwent an “examination of linguistic patterns and their functions” (Bellard-Thomson, 2011, p. 55) and considered the “‘effects’ of such patterning” (p. 55) that were a result of authors’ uses of the target voice features—with the anticipated learner outcome of improving authorial voice in response to TOEFL iBT independent writing task prompts. In this manner, the primary function of the stylistic analyses was to engage learners—through the analysis of literary texts—in ways that “explore and exploit the resource of language in original ways” (Lin, 2010. p. 193). In the context of this study, “original” is interpreted as learners employing the target language features to meet their writing goals.

To operationalize these “examinations of linguistic patterns” I provided participants with stylistics-based worksheets (Appendix G) that were employed to analyze literary texts. The worksheets, adapted from Lin (2010), were used successfully during the pilot phase of this study. The worksheets afforded participants opportunities to identify the target voice



features and afforded participants opportunities to describe the semantic and/or pragmatic function of these features in a given text (see Section 3.3 for a detailed discussion of the content of the worksheets). Below I discuss how the stylistic analyses were completed.

#### 4.5 Classroom activities

I describe week one separately, as many of the activities were introductory or explanatory in nature. These activities differed markedly in their goals from the activities employed during weeks two and three (see too Table 4.2 for descriptions of the classroom activities across the intervention phase of this study).

##### 4.5.1 Week one (Days one through four)

The first of the six instructional stages informing STI involves presenting learners with the “what and how something is to be done” (Lantolf & Thorne, 2006, p. 304). This stage included explicitly stating the aim of the learning task (in this case, improving the quality of voice in essay responses to the TOEFL iBT independent writing task) and the steps required to achieve the goal. The first half of Day 1 of the intervention was devoted to guiding participants through stage one. The second half of Day 1 introduced stage two of STI.

The second stage of STI introduced an appropriate OBA (orienting basis of an action) to provide learners with a clear, visual step-by-step outline of how to achieve the target. Galperin suggested that this outline be presented as an orienting chart or diagram (Lantolf & Poehner, 2014) “that gives the conceptual properties of the action a material form” (Johnson, 2008, p. 73). Day 1 and Day 3 introduced participants to the OBA (see Appendix L) with a

particular emphasis on introducing the scientific concepts of the five rhetorical features targeted in this study. (Days 3, 7, 9, and 11 also provided participants with opportunities to employ the OBA.) Days 2, 3, and 4 also introduced stylistics as a mediating tool for exploring language, and model stylistic analyses were also conducted in class during these three days. While introducing stylistics and stylistic analyses participants were provided with opportunities to develop a comfortable understanding of the processes and techniques involved in performing a stylistic analysis of a literary text.

Other activities during week one centered on different aspects of the TOEFL independent writing task associated with voice. These activities included watching a video produced by ETS that outlines how to successfully navigate the independent writing task (discussion questions I posed in class regarding this video were framed by the concept of voice in academic writing), examining exemplary model and poor sample essays provided online by ETS (again, by employing a lens framed by voice in academic writing), as well as working through the scoring rubric for the TOEFL independent writing task. (Table 4.2 describes these activities in more detail and provides the aim and rationale for introducing these activities.)

#### 4.5.2 Week two and three (Days five through twelve)

Weeks two and three afforded participants with opportunities to develop stages three through six of STI through mediating artifacts and classroom activities. For example, stage three was characterized by employing didactic models, including diagrams, tables, or illustrations to mediate concept development. Although the ETS video employed on Day 2 was a mediating tool associated with stage three, the central didactic model (Visual Aid II—

Appendix M) for stage three was introduced on Day 5. Visual Aid II was a handout that provided participants with a different visual understanding of how to navigate the function and use of the five target features. Moreover, unlike the online video, Visual Aid II was an easily accessible tool that participants could easily utilize in class.

Stages four and five were characterized by different modes of languaging the scientific concept to aid with internalizing the target features (Swain et al., 2009). Appendix G provides an example of the worksheets that were employed almost daily during the intervention to facilitate these stages. Other exercises (see Table 4.2) also provided opportunities for languaging, including exercises that elicited collaborative peer feedback regarding the quality of voice in participant essays, opportunities for participants to rework their essays (Days 3, 5, 7, 9, and 11), and exercises that asked participants to analyze model essays (individually and in groups) with the conceptual units in mind (Days 3, 7, and 9). Table 4.2 outlines the instructional sequence of weeks two and three and provides details regarding the procedures, aims, rationale, and activities employed to foster development.

Weeks two and three also afforded participants opportunities to demonstrate having realized stage six of STI—the complete scientific concept development of the five rhetorical features. These opportunities were presented on Days 4, 6, 8, 10, and 12 of the intervention phase (as well as during the delayed posttest). On these days participants responded to an ETS-produced essay prompt designed for the TOEFL iBT independent writing task. As noted in Chapter 3, these essays were completed in a computer lab and were the main component of the data I collected for this study.

#### 4.6 Feedback during the instructional phase

I also provided feedback on the written compositions produced by the participants throughout the intervention phase. The type of feedback I provided combined an interactionist approach to dynamic assessment (this approach focuses on interactions between the instructor and the learner to mediate learning, as opposed to an interventionist approach to dynamic assessment that concentrates on standardized forms of assistance), as outlined in Lantolf and Thorne (2006, 2007), with the 13-point regulatory scale (shown in Figure 4.1) created by Aljaafreh and Lantolf (1994).

An interactionist approach to dynamic assessment framed the content of the feedback I provided in the margins of participants' essays. These comments allowed me to direct participants' learning with the aim of facilitating their concept development. I employed Aljaafreh and Lantolf's (1994) 13-point regulatory scale as a tool for conducting consistent and self-monitored feedback. (After using this scale during the pilot study, I decided to begin my feedback at point three or point five on the scale, depending on the writing error; this decision was made primarily because of time constraints in the classroom.) Although initially designed for tutors and their students, I employed this regulatory scale in the classroom as a tool for providing individual feedback for each learner. The scale allowed me to adjust my feedback—almost synchronously as I moved from participant to participant within the classroom—along a scale from implicit to explicit feedback. This monitoring allowed me to provide graduated assistance (Aljaafreh & Lantolf, 1994) toward concept development and afforded participants opportunities to work through their errors systematically. Although the example interaction between tutor and learner employed in Aljaafreh and Lantolf's 13-point regulatory scale is focused on language forms, I also provided feedback on language

functions pertaining to voice. For example, on Tom's Day 3 written composition I wrote the following comment in the margin: "Do you think using more 'I' in your essay could make your position or your opinion stronger?"

0. Tutor asks the learner to read, find the errors, and correct them independently, prior to the tutorial.
1. Construction of a "collaborative frame" prompted by the presence of the tutor as a potential dialogic partner.
2. Prompted or focused reading of the sentence that contains the error by the learner or the tutor.
3. Tutor indicates that something may be wrong in a segment (e.g., sentence, clause, line)- "Is there anything wrong in this sentence?"
4. Tutor rejects unsuccessful attempts at recognizing the error.
5. Tutor narrows down the location of the error (e.g., tutor repeats or points to the specific segment which contains the error).
6. Tutor indicates the nature of the error, but does not identify the error (e.g., "There is something wrong with the tense marking here").
7. Tutor identifies the error ("You can't use an auxiliary here").
8. Tutor rejects learner's unsuccessful attempts at correcting the error.
9. Tutor provides clues to help the learner arrive at the correct form (e.g., "It is not really past but some thing that is still going on").
10. Tutor provides the correct form.
11. Tutor provides some explanation for use of the correct form.
12. Tutor provides examples of the correct pattern when other forms of help fail to produce an appropriate responsive action.

*Figure 4.1* Regulatory scale—implicit (strategic) to explicit. From "Negative feedback as regulation and second language learning in the zone of proximal development," by A. Aljaafreh and J. Lantolf, 1994, *The Modern Language Journal*, p. 475. Copyright 1994 by Wiley-Blackwell. Reprinted with permission.

#### 4.7 Narrative account of teaching-learning interactions

To document firsthand the learning and teaching context, this section provides a narrative account of the daily interactions in the classroom. The class convened around one conference table where students formed a semi-circle facing the front of the class, and on most days a slide-based presentation was projected from the outset. After informal greetings, a class typically began with a review of one of two items: the previous day's stylistic analysis or the previous day's writing samples that I had collected as part of my dataset. Below I discuss how these items were reviewed in class.

First, regarding the previous day's stylistic analysis, the class would first reread the text that was analyzed. I would then return students' original analyses that were conducted on the stylistic worksheets (Appendix G) to them (I had already made copies of the analyses for my files). After students reviewed their analyses to refresh their memories, I would lead students to discuss their findings. This enabled a class-wide discussion, rather than the small-group discussions that informed the analyses on the previous day. To initiate the discussion, we reviewed the first section of the analysis that asked students to identify and label the rhetorical features that they had already identified. Herein students would participate freely and lead the review by noting what features they had highlighted. Other students would then interject, reporting other instances of rhetorical features they found, or students would ask each other (or more often, they would ask me) if they had labelled a particular feature correctly. I acted primarily as a facilitator to the discussion and provided clarity when different interpretations of rhetorical features were offered.

When the class (myself included) was satisfied that we had correctly identified and labelled all of the target rhetorical features, the interaction shifted focus. At this next stage

we discussed how particular discourse features influenced the text in different ways. While the students usually engaged each other, on occasion I was required to encourage students to participate. To do so I highlighted sample answers (anonymously) from students' worksheets that I thought were particularly insightful. This approach encouraged students to add their own opinions to the discussion. Moreover, and where I thought necessary, I also asked students to probe more deeply into their responses. For example, I would present the class with partial interpretations of why a particular rhetorical feature was used, or I would ask students which of two or three possible interpretations might work best. In these instances students were encouraged to defend their interpretations by referencing the text. Students were provided with time to think about a possible response, so there were often moments of silence during these discussions (one to two minutes, typically) when students were reflecting on a possible response. During this silence I could see students rereading the text and making short notes.

As the intervention progressed these discussions became more stimulating and insightful, and students became more engaged and seemingly more confident with their work: Students were better able to defend their interpretations of a given text based on the discourse features we had analyzed. Consequently, my role interacting in or facilitating discussions diminished over time.

Second, when I started a class with a review of the latest writing prompt to which students had responded, I began by distributing to students a hard copy of their essays. After students reread their essays to refresh their memories, they were asked to focus on a slide presentation I had prepared. The slides would either document exemplary or problematic patterns in students' writing that I considered worth addressing, or the slides would provide

review material regarding essay-writing strategies, or both. For example, in one class I presented how the location of the first (and subsequent) direct-reader reference in an essay could influence reader's engagement levels. At other times I addressed questions regarding the topic I presented. These topics were delivered in a lecture format that engaged the students with questions or brief discussion tasks. These brief lectures ended with a short task that was based on the lecture material. For example, students exchanged papers with a peer and were asked to comment in the margins on discourse features or other elements of writing that paired with the content of the lecture. When the essays were returned, students briefly engaged in peer discussions about the comments they had just received.

When these discussions concluded I distributed to students another copy of their essays. This copy included comments in the margins that I had made, and this feedback was framed on the 13-point regulatory scale (Figure 4.1) noted in Section 4.6. Students were then given time to revise sections of their essays based on my comments, on the comments of their peers, or based on comments I had made at the beginning of class. At this time I circulated around the class, making a point to visit with each student at least once to provide individual attention and to address any writing concerns particular to individual learners.

Barring Day 1 and Day 7, all of the classes during the intervention phase then concluded with the same final task: a stylistic analysis of a literary text, for which I provided background information on multiple slides. This presentation usually included biographical data about the author or poet, a photograph, and a short oral narrative that I relayed regarding the personal lives or regarding the accomplishments of the writers. For example, prior to analyzing Margaret Atwood's poem "This is a photograph of me" I highlighted that Atwood is a famous and prolific Canadian author who has won numerous literary prizes, and that her



works are translated into many languages, including Japanese. This biographical information was provided to stimulate learners' interest in the text and to showcase the authors as exemplary users of the target voice features that are, thus, worth emulating.

After distributing the worksheets and a copy of the poem, I also introduced some vocabulary that appeared in the literary texts (based on my experiences with Japanese learners at this proficiency level, I selected vocabulary that I inferred would be unfamiliar to this group of students). Students then worked individually through the first three questions on the stylistic analysis worksheet. During the first few days of the intervention students would occasionally ask questions that clarified for them the demands of the task, and I would answer these questions on an individual basis. Many students also turned to their mobile phones or electronic dictionaries for help with understanding vocabulary that I had not reviewed. Twenty minutes was allocated for this phase of the stylistic analysis. Following this initial phase, students then shared their findings in small groups for ten minutes. During this time I set down recording devices and did not interfere with the discussions. Class ended when these discussions were completed, and I collected the stylistic analyses to be included in my dataset.

#### **4.8 Summary**

This chapter documented the instructional context for this study. This documentation included an outline of the teaching qualifications of the researcher and detailed profiles of the ten participants. The six stages that comprise STI were also reviewed, explaining how each stage was operationalized within the intervention phase of this study. This chapter also outlined how stylistics was employed in this study and provided a detailed report of the instructional content and activities that informed the intervention stage of this study. The

chapter concluded with a narrative account of the daily teaching and learning interactions that occurred during the intervention.

## Chapter 5

### Findings about Stylistics, Authorial Voice Development, and CBI

In this chapter I present the findings from the statistical analyses performed on the data I collected for this study alongside relevant verbal data from the post-intervention interviews. The findings are arranged below by each research question (but note that RQ1 is composed of two parts, which I address independently).

**5.1 RQ1** *Is there an increase, over the intervention period, in participants' (a) overall quality of writing, and (b) overall quality of voice in writing?*

Table 5.1 documents the means and standard deviations for change over time in the overall quality of writing and the overall quality of voice in participants' writing. (Appendix N provides graphical representations of these changes.)

Table 5.1  
*Summary of Means and Standard Deviations for Writing Scores and Voice Scores Over Time for Participants in the Intervention (n = 7)*

Variable	Baseline		Day 4		Day 6		Day 8		Day10		Posttest		Delayed posttest	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Writing	2.37	0.58	2.47	0.53	2.6	0.38	2.89	0.43	2.96	0.55	3.16	0.35	3.17	0.39
Voice	2.37	0.59	2.93	0.77	2.56	0.46	2.91	0.55	2.8	0.75	3.4	0.56	3.11	0.6

To answer RQ1 Friedman tests were used to detect any statistically significant differences over time across the writing samples. Statistically significant differences over time were detected regarding the overall quality of writing ( $\chi^2(6) = 20.86, p = .002$ , with a

large effect size,  $W = .5$ ) and the overall quality of voice in writing ( $\chi^2(6) = 21.37, p = .002$ , with a large effect size,  $W = .51$ ).

Pairwise comparisons were then conducted using Wilcoxon signed-rank tests, with a Bonferroni correction applied. Table 5.2 shows the results of the post hoc tests on the overall quality of writing. The post hoc analyses were unable to detect significant changes. This is unsurprising given that the Bonferroni correction is conservative (Cabin & Mitchell, 2000; Field, 2009; Larson-Hall, 2010; Perneger, 1998). The effect sizes ( $r$ ) reported in Table 5.2 show four pairwise comparisons that demonstrate a large effect size and twelve pairwise comparisons that demonstrate a very large effect size. The largest of these effect sizes appear in comparisons between the written compositions collected on Day 6 and the posttest ( $z = -2.37, p = .018, r = -.9$ ) and Day 6 and the delayed posttest ( $z = -2.37, p = .018, r = -.9$ ).

Table 5.3 shows the results of the post hoc tests on the overall quality of voice in writing. Herein the analyses revealed statistically significant changes with pairwise comparisons between the written compositions collected on Day 10 of the intervention and the posttest (Day 12),  $z = -2.41, p = .016$  (with  $p$  set at  $.025$  based on a Bonferroni correction,  $.05/2$ ),  $r = .91$ , and between the written compositions collected during the posttest and the delayed posttest (four weeks after the posttest),  $z = -2.03, p = .042$  (with  $p$  set at  $.05$  based on a Bonferroni correction,  $.05/1$ ),  $r = .77$ . The effect sizes ( $r$ ) reported in Table 5.3 show that there are four pairwise comparisons that demonstrate a large effect size and seven pairwise comparisons that demonstrate a very large effect size. The largest of these effect sizes appear with the written compositions collected on Day 6 and Day 8 ( $z = -2.4, p = .016, r = -.91$ ) and Day 10 and the posttest ( $z = -2.41, p = .016, r = -.91$ ).

To supplement these findings, and to provide further insight into the changes noted

Table 5.2  
 Follow-up Pair-wise Tests Regarding Differences Over Time in the Overall Quality of Writing ( $n = 7$ )

	Day 4			Day 6			Day 8			Day 10			Posttest			Delayed posttest			
	* $p <^a$	$z^b$	Sig. <sup>c</sup>	$r^d$	$z$	Sig.	$r$	$z$	Sig.	$r$	$z$	Sig.	$r$	$z$	Sig.	$r$	$z$	Sig.	$r$
			$d^e$			$d$			$d$			$d$			$d$			$d$	
Baseline	.0083 (.05/6)	-1.38	.168	-.52 -1.22	-1.61	.107	-.61 -1.54	-2.21	.027	-.84 -3.1	-2.12	.034	-.8 -2.67	-2.2	.028	-.83 -2.98	-2.2	.028	-.83 -2.98
Day 4	.01 (.05/5)				-1.19	.233	-.45 -1.01	-2.12	.034	-.8 -2.67	-1.78	.075	-.67 -1.81	-2.22	.027	-.84 -3.1	-2.21	.027	-.84 -3.1
Day 6	.0125 (.05/4)							-2.22	.027	-.84 -3.1	-1.78	.075	-.67 -1.81	-2.37	.018	-.9 -4.13	-2.37	.018	-.9 -4.13
Day 8	.0167 (.05/3)										-0.81	.416	-.31 -.65	-2.21	.027	-.84 -3.1	-2.21	.027	-.84 -3.1
Day 10	.025 (.05/2)													-1.19	.233	-.45 -1.01	-1.19	.233	-.45 -1.01
Posttest	.05 (.05/1)																-0.45	.655	-.17 -.35

<sup>a</sup> based on a Bonferroni correction. <sup>b</sup> Wilcoxon signed-rank test. <sup>c</sup> Asymp. sig. (2-tailed). <sup>d</sup> Effect size ( $r = z/\sqrt{N}$ ). <sup>e</sup> Standardized effect size ( $d = 2r \div \sqrt{1 - r^2}$ ).

Table 5.3  
 Follow-up Pair-wise Tests Regarding Differences Over Time in the Overall Quality of Voice in Writing ( $n = 7$ )

	Day 4			Day 6			Day 8			Day 10			Posttest			Delayed posttest			
	* $p <^a$	$z^b$	Sig. <sup>c</sup>	$r^d$	$z$	Sig.	$r$	$z$	Sig.	$r$	$z$	Sig.	$r$	$z$	Sig.	$r$	$z$	Sig.	$r$
			$d^e$			$d$			$d$			$d$			$d$			$d$	
Baseline	.0083 (.05/6)	-1.69	.09	-.64	-0.51	.61	-.19	-1.49	.138	-.56	-1.27	.204	-.48	-2.2	.28	-.83	-2.2	.28	-.83
			-1.67			-.39			-1.35			-1.09			-2.98			-2.98	
Day 4	.01 (.05/5)				-1.07	.285	-.4	-0.17	.863	-.06	-0.74	.458	-.28	-1.69	.092	-.64	-0.64	.525	-.24
						-.87			-.12			-.58			-1.67			-.5	
Day 6	.0125 (.05/4)							-2.4	.016	-.91	-0.95	.343	-.36	-2.38	.017	-.09	-2.21	.027	-.84
												-4.39			-.18			-3.1	
Day 8	.0167 (.05/3)										-0.14	.892	-.05	-2.23	.026	-.84	-1.18	.239	-.45
													-1		-3.1			-1.01	
Day 10	.025 (.05/2)													-2.41	.016*	-.91	-1.45	.147	-.55
															-4.39			-1.32	
Posttest	.05 (.05/1)																-2.03	.042**	-.77
																			-2.41

<sup>a</sup> based on a Bonferroni correction. <sup>b</sup> Wilcoxon signed-rank test. <sup>c</sup> Asymp. sig. (2-tailed). <sup>d</sup> Effect size ( $r = z/\sqrt{N}$ ). <sup>e</sup> Standardized effect size ( $d = 2r \div \sqrt{1 - r^2}$ ). \*  $p < .025$ , \*\*  $p < .05$

thus far in learners' overall quality of writing and in learners' overall quality of authorial voice, in the subsequent paragraphs I document results derived from an analysis of the verbal data. Table 5.4 summarizes participants' responses to interview question I.A.2 (see Appendix H) regarding their posttest written compositions. In this question participants were asked to comment on classroom activities that may have influenced their uses of the rhetorical feature in question.

**Table 5.4**  
*Tally of Responses Regarding Mediating Tools and Rhetorical Feature Use (n = 7)*

	Mediating tools			
	QI.A.2: Is there anything that we studied in class or discussed in class that may have influenced this [rhetorical feature] choice?			
	Stylistics	Sample essays	Other	No recollection
Booster	2	3	1 (OBA <sup>a</sup> )	1
Hedge	2	2	2 (OBA)	1
Authorial self-mention	4	2	1 (Writing sample feedback)	0
Attitude marker	2	2	0	3
Direct reader reference <sup>b</sup>	4	0	0	1

<sup>a</sup> Orienting basis of action.

<sup>b</sup> There were only five responses here as two participants did not use direct reader references in their posttest.

Table 5.4 highlights two classroom-based mediating tools (stylistic analyses and sample essay analyses) as the potential sources for the statistically significant change over time in writing and the statistically significant change over time in voice scores noted in response to RQ1. Below I elaborate on these two sources of change by first documenting

verbal data related to change in voice scores and then by noting verbal data related to change in writing scores.

As a tool for mediating voice development, participants commented that stylistics influenced uptake across all five of the target rhetorical features. The following comments speak to this influence (Appendix U describes the transcription conventions):

- “Through analysis of poem I learned about hedge. Like when I wanna say my opinion but not (+) it doesn’t fit for everyone, it’s better to use hedge” (Tom, post-intervention interview, September 29, 2013);
- “Because (+) there are many [boosters] in short stories so I think it is good for essay” (Am, post-intervention interview, September 29, 2013);
- “Through the classes to analyze the poem and small novels [short stories] so I (+) understood which stylistic feature will be a good for the essay” (Pi, post-intervention interview, September 29, 2013); and
- “I think stylistics is interesting. There is not so many chance to know how does this word works in the writing. So yea (+) yea, it is very useful for me” (Angela, post-intervention interview, September 29, 2013).

More specifically, stylistics was cited as influential in guiding students through the concept development of direct reader reference (stylistics is prevalent here as a mediating tool despite two participants not using direct reader references in their posttests) and authorial self-



mentions. The following comments exhibit the influence of stylistics in developing proficiency of these two rhetorical features:

- “Many authors use direct reader reference. When I read that sentences it was very strong. (+) They make me have to think about it [the sentence topic] ((Right.)) So I want the reader to think about this [online purchases—the topic of the paragraph in question] [...] Many authors use direct reader reference very effectively, so I want to use too” (Maru, post-intervention interview, September 29, 2013);
- “I use many ‘I’ (+) like use in [song] lyrics and short stories” (Am, post-intervention interview, September 29, 2013); and
- “Many times we circled many authorial self-mention. That’s why I, like, did this [employed a series of authorial self-mention]. ((What were you circling in class?)) Oh (+) those times with poems and stories, like ((Right.)), we had to find them [authorial self-mention]” (Tom, post-intervention interview, September 29, 2013).

These comments address participants’ perspectives on the effectiveness of stylistics as a mediating tool for developing the scientific concepts of direct reader reference and authorial self-mention.

The participants also observed that examining sample essays informed their concept development of the target voice features. This influence was strongest regarding the concept development of boosters. The following comments describe this influence:

- “I saw ‘strongly’ [the booster in question] on the screen [sample essay displayed in class via a PowerPoint presentation] so I thought it’s good to use” (Angela, post-intervention interview, September 29, 2013).
- “I saw many examples [of boosters]. ((What kind of examples?)) Essay examples” (Pi, post-intervention interview, September 29, 2013).

Of the two dominant influences noted in Table 5.4 (stylistic analyses and sample essay analyses), stylistics was the dominant mediating tool across all of the target features, with a pronounced influence on the use of direct reader reference and authorial self-mention.

Interview Question II.B.3 (Appendix H) asked participants to comment on stylistics as a tool for mediating voice development. Table 5.5 summarizes their responses. These comments deal with the effectiveness of stylistics for mediating authorial voice development and also address microgenetic dimensions of change in learners’ writing behaviour. For example, these comments apply to the source, variability, and breadth of learners’ microgenetic processes. First, the comments are consistent in indicating that stylistics was the primary source of change in learners’ writing behaviour. Second, this consistent reference to stylistics as informing change in learners’ writing emphasizes a lack of variability between participants in documenting the source of their learning. Third, with the exception of Angela and Anne, the comments in Table 5.5 speak to learners’ awareness of the transferability of the target learning features from literary texts to their own academic writing. Such awareness may contribute to improvements in their overall quality of writing and overall quality of authorial voice, thus emphasizing learners’ breadth of learning.

Table 5.5

*Summary of Comments Regarding the Role of Stylistics as a Mediating Tool for the Concept Development of the Target Rhetorical Features (n = 7)*

Perspectives on stylistics as a mediating tool for developing authorial voice	
Q11.3: In general, what do you think about using stylistics to learn about voice in academic writing?	
Maru	Yea... it was amazing for me because many poem and some stories are very short. If I read very thick book I think they don't [use] so much voice features (+) ((Okay.)) because they can explain detail but in song or poem they are condense their idea. They [authors and poets] have to condense their ideas so need to explain very short but effective expression so we analyze these poems (+) so we can get, 'Ah! How to use these expressions or [voice] features'. So this essay is also short so it's (+) it's not similar but it's same [length, i.e., short] as poem so to use these kind of words [lexical voice features] or expressions is good because effective.
Tom	I think it's [stylistics] a really good idea because they [the analyses] influence me to use hedge and booster and each [feature]. Each author has a habit of writing. So when I saw many types of habit I could see my habit as well. ((Can you explain more?)) In my case there are not enough voices [voice features] but author use very well those voices [voice features] to say their opinion or feeling so I thought, 'Oh. I have to steal this skill'. But you know three weeks are not enough to improve a lot because authors are professionals so sometimes it's really too difficult to know why authors use those [voice features] so I couldn't analyze well in short time. So if I analyze more and more I think I can analyze my essay better as well to improve.
Anny	Things like booster or hedge, or how to explain our feeling (+) stylistics is very good to explain our opinion. If the question [essay prompt] is asking opinion, it [stylistics] is really helpful. (+) This time I studied more casual like poems. It's easy to understand. I can feel if it's [author's voice] strong or not. (+) Studying poetry helpful because easier to see [voice] features over newspaper articles [or other non-literary texts].
Angela	I think it's good to know how famous authors or writers use language.
Pi	If I just learned or know the (+) like hedge (+) just the word it's not useful to write essay because I don't have examples. ((Okay.)) Examples is very important to use those things [voice features]. ((Which type of examples did you prefer: model essays or stylistics?)) Um (+) I prefer the poem and story because you have a lot of variety (+) not so general. There was lots of examples of using the [voice] features for us to look at. [...] To study a variety of poem and short story is good for (+) good to write essay because we can see the examples through a story or poems.
Am	Songs and short stories helpful but I think poem is too difficult for me. [...] They [songs and short stories] are easy to understand how writer using language so it's good for me too for writing.
Anne	The poem is example is helpful for us but if I don't know very well about the author [the author's biography] I cannot interested about the poem because if I know just only the poem (+) the poem is more confusing for me but if I can know about the author before studying the poem, the poem more (+) more interesting. I can understand the author's decision with writing.

Regarding changes in participants' overall quality of writing, learners in the intervention group attributed this improvement to an increased awareness (triggered, participants remarked, by stylistics) of how to employ the target rhetorical features. The following are examples of participants' commentary on this point:

- “I know these words [voice features] but I can't explain (+) I thought I just tell my idea but that is like mechanic (+) like a robot. So I learned this voice [features]. This is (+) very supports my ideas and my opinion and my position so it's perfectly positive effect for me. ((Can you provide an example, maybe?)) Um (+) if I didn't know the voice or the stylistics (+) I will write 'I also have an experience to live in a small town when I was a university student.' ((So you would have skipped the word '*fortunately*'?)) Yes. My introductions are perfectly different now” (Maru, post-intervention interview, September 29, 2013);
- “First, um (+) I didn't know about voices [voice features], especially in stories, so I was always thinking why my essay or my sentence in English are really simple. ((Okay.)) So your classes on voice gave me hint to improve my essay. ((Can you think of an example?)) After your lectures I always think more (+) what more can I add to my sentence so I think hedge and

boosters and influence to my essay when I say opinion or feeling” (Anny, post-intervention interview, September 29, 2013); and

- “If I use these tips [regarding voice features] I can write better” (Angela, post-intervention interview, September 29, 2013).

These last two comments by Angela and Anny were also paired with negative appraisals of their writing as a result of the intervention:

- “It’s difficult to make ideas and use voice features at the same time” (Angela, post-intervention interview, September 29, 2013); and
- “I came to be able to not finish my essay. I was confused. ((What were you confused about?)) I tried to think about more things. Before I write essay just quickly and finish but now I want to make my essay more good or nice so I think before I begin to write. Then I don’t have enough time” (Anny, post-intervention interview, September 29, 2013).

These comments also reveal awareness of control over how to employ (and the effectiveness of employing) the target rhetorical features, across all of the participants.

The verbal data presented in the preceding paragraph highlights learners’ awareness of how to employ the target discourse features; other verbal data consistently emphasized awareness of the dialogic space that learners participate in as essay writers. For example, these learners expressed a conscious desire to engage readers with their writing and expressed a conscious desire to mark their authorial stances clearly on a particular topic. In

Section 5.3 I document and extrapolate on consistent commentary regarding the use of direct reader references and authorial self-mention for this end.

Other verbal data also address the effectiveness of explicit instruction (a hallmark of CBI-informed pedagogies) for developing learners' writing proficiency. Specifically, the findings noted here concern explicit CBI-related tasks and explicit CBI-related mediating tools employed in this study. (The verbal data presented in the subsequent paragraphs were collected across the interviews, as opposed to data collected in response to a particular question, as presented in Tables 5.4 and 5.5.)

All of the participants remarked that stylistics (presented explicitly as a mediating tool for developing proficiency in the target voice features) was an effective tool for mediating the concept development of the target rhetorical features. This observation was recorded on 17 occasions (Table 5.5 provides sample commentary in this regard). As a form of explicit instruction that addressed strong and weak writing samples, six participants (except Tom) commented on eight occasions that sample essays functioned as an effective mediating tool. Five participants (excluding Anny and Am) also noted that explicit comments made by the instructor on effective essay writing informed decisions about their writing. The following are examples of participant commentary in this regard:

- “Um (+) I use a lot ‘I’ or ‘my’ because I learned in your talking it’s (+) more stronger to use ‘I’ (+) like my experience. ((Right.)) Because um (+) if I didn’t have experience about this it’s a little weak point but to use ‘my’ (+) to write my experience (+) I can tell the reader my strong opinion” (Pi, post-intervention interview, September 29, 2013);

- “In the TOEFL 4 [Day 10] (+) you said [in TOEFL 4 essay feedback] there is only 1 authorial self-mention so (+) I think I have to use more voice [features]” (Angela, post-intervention interview, September 29, 2013);
- “Booster can make my essay strong, so (+) you told me to use many boosters. It makes my essay look confidently, so I did” (Maru, post-intervention interview, September 29, 2013); and
- “Before posttest (+) I couldn't develop each body [paragraph]. So (+) this time, introduction and first paragraph is smaller than before but I could write second and third [body] paragraph so I think it was better than before. ((Okay. Was there anything we studied in class that, perhaps, influenced this change?)) First baseline test (+) at that time I didn't know (+) how to write with voice but you told me many things like voice features is very effective so in the second [body] paragraph, ‘*When I wanted a personal computer in a small city, I couldn't find the computer that I wanted*’ (+) actually if I didn't know [about] voice features I finished here. ((Right.)) But I want to use direct reader reference so, ‘*You might think that we can buy almost all products by online.*’ This idea [directly addressing the reader] is from direct reader reference so if I didn't know about the [voice feature]

idea I couldn't write [this]" (Maru, post-intervention interview, September 29, 2013).

Less marked comments also speak to the effectiveness of the OBA as an explicit mediating tool. Three intervention group participants (Maru, Tom, and Angela) commented positively on the OBA. For example, Tom noted the following: "I think when I was writing I was reminded of your graph. ((Which one?)) It had arrows or ((The OBA?)) Yea. Yea. Yea. ((Okay.)) It helps me a lot. Like to choose booster or hedge or something like that" (post-intervention interview, September 29, 2013). Angela reported the following regarding her use of the OBA: "The OBA can help my essay directly. ((How?)) There are concrete advices for me. I can follow the OBA. [...] Whenever you showed us model essay I realized how important it [the OBA] is" (post-intervention interview, September 29, 2013). The remaining participants (Anny, Pi, Am, and Anne) expressed less favourable commentary regarding the OBA. Anny found the OBA distracting, noting that she was unable to transform the examples into her own writing. Pi was unable to recall the OBA while responding to essay prompts, noting that there were insufficient examples to guide her. Am expressed the same sentiment regarding a lack of extensive examples. Anne thought the OBA was unhelpful and chose not to reference it during the course.

#### 5.1.1 Summary of findings for RQ1 (part 1)

Using Friedman tests, statistically significant differences over time were detected regarding students' overall quality of writing and the overall quality of voice in writing.



Except for results from two pairwise comparisons regarding the overall quality of voice in writing (Day 10 and the posttest, and the posttest and the delayed posttest), post hoc analyses were unable to detect statistically significant findings for the overall quality of writing that the Friedman test had highlighted. For this reason, and to offset the risk of a Type II error that emerges from using the Bonferroni correction, effect sizes were emphasized in reporting the post hoc analyses.

The verbal data reported in Section 5.1 provide insight into the changes over time evident in learners' writing and learners' quality of voice in writing, with particular respect to (a) the influence of stylistic analyses and sample essay analyses on developing writing and authorial voice over the intervention period (with, regarding voice development, emphasis on direct reader references and authorial self-mention), (b) learners' awareness of their role as authors in a dialogic space, and (c) the role of explicit instruction and the influence of the OBA in a CBI-informed learning context. From a microgenetic perspective, the verbal data reported in Section 5.1 emphasized (a) the role of stylistics as the primary source of change informing learners' writing behaviour, (b) the lack of variability in indicating stylistics as the main source of change across the participant group, and (c) the role that learners' awareness contributed to the breadth of learning across the participant group, with the exception of Angela and Anne.

5.1.2 RQ1 (part two) *Is any increase in the quality of writing and the quality of voice sustained after the intervention?*

To answer the second part of RQ1, I conducted a Wilcoxon signed-rank test that compared essay scores and voice scores collected from the posttest and, four weeks later, the

delayed posttest. Rank-orders were the basis for these analyses. Table 5.6 shows the means, standard deviations, and medians regarding change in the overall quality of writing and the overall quality of voice as well as the results from the Wilcoxon signed-rank test.

Table 5.6

*Summary of Means, Standard Deviations, and Wilcoxon Signed-rank Test for Writing Scores and Voice Scores Sustained Over Time (n = 7)*

	Posttest			Delayed posttest			z	r <sup>a</sup>	Sig. <sup>b</sup>	d <sup>c</sup>
	M	SD	Mdn	M	SD	Mdn				
Writing	3.16	0.35	3.2	3.17	0.39	3	-0.45	-.17	.66	-.35
Voice	3.40	0.56	3.3	3.11	0.6	3	-2.03	-.77	.04*	-2.41

<sup>a</sup> Effect size ( $r = z/\sqrt{N}$ ). <sup>b</sup> Asymp. sig. (2-tailed). <sup>c</sup> Standardized effect size ( $d = 2r \div \sqrt{1 - r^2}$ ).  
\*  $p < .05$

Regarding the overall quality of writing, mean essay scores in the participant group demonstrated no change over time from the posttest to the delayed posttest. There was a small effect size. This result indicates that increases in the quality of writing observed during the intervention phase of this study were sustained over time. Regarding the overall quality of voice, the mean voice scores in the participant group decreased over time. This decrease was a statistically significant change from the posttest to the delayed posttest. There was a large effect size. This result indicates that increases in the quality of voice observed during the intervention phase of this study were not sustained over time, and from a microgenetic perspective, the rate of change concerning authorial voice development was gradual (and ongoing), rather than abrupt. Appendix O provides a graphical representation of the changes noted here.

The verbal data provides only one insight into why changes in learners' overall quality of voice were not sustained:

But you know three weeks [the length of the intervention] are not enough to improve a lot because authors are professionals so sometimes it's really too difficult to know why authors use those [voice features] so I couldn't analyze well in short time. So if I analyze more and more I think I can analyze my essay better as well to improve. (Tom, post-intervention interview, September 29, 2013).

Tom's comments suggest that a longer intervention period may have provided greater opportunities for internalizing the target voice features.

### 5.1.3 Summary of findings for RQ1 (part two)

The second part of RQ1 sought to examine whether any increases in the quality of writing and the quality of voice in writing were sustained during the four-week period that separated the posttest from the delayed posttest. The increases in the quality of writing enjoyed by the intervention group were sustained over this period of time. In contrast, the overall quality of voice in writing was not sustained over time, indicating from a microgenetic perspective only a gradual and as yet not fully realized change in authorial voice development. One participant suggested that a longer period of study may have afforded her a better opportunity to internalize the target features.

**5.2 RQ2** *What is the relationship between the frequency of use of each target rhetorical feature of voice, and (a) the overall quality of writing and (b) the overall quality of voice?*

To answer RQ2 I first calculated the means and standard deviations regarding the frequency of use of each target rhetorical feature of voice in students' written compositions.

Table 5.7 documents these data.

**Table 5.7**  
*Summary of Means and Standard Deviations for Frequency of Use of Rhetorical Features Over Time Across 49 Essays (n = 7)*

Variable	Baseline <sup>a</sup>		Day 4		Day 6		Day 8		Day10		Posttest		Delayed posttest	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Hedge	1.43	0.52	1.14	1.21	3	2.08	2	1.83	4.14	3.29	2.43	0.79	1.86	0.9
Booster	5.91	2.64	7.86	3.85	6.86	1.35	7.57	3.31	5.43	3.51	9.57	3.87	6.57	3.69
Attitude marker	3.57	1.6	2.86	2.04	1.86	1.21	2.57	1.62	3.14	1.21	2.29	1.38	1.57	1.13
Authorial self-mention	10.76	4.44	17.43	8	7	4.65	4.71	3.45	10.43	5.91	12.43	7.16	11.29	7.25
Direct reader reference	5.38	5.84	7.29	5.06	6.57	4.76	9.29	4.54	3.43	2.99	6.86	8.17	10.29	5.94

<sup>a</sup> Baseline scores are derived from the means of 3 baseline, essay-writing tasks.

Table 5.7 demonstrates that boosters, authorial-self mentions, and direct reader references occurred more frequently across the dataset than did hedges and attitude markers.

After calculating the means and standard deviations, I then employed Spearman's rho. Correlations were examined by analyzing the group mean usage of each rhetorical feature and the group mean essay scores from the TOEFL iBT independent writing tasks, and then the group mean usage of each rhetorical feature and the group mean voice scores

collected from the same writing tasks. Table 5.8 documents the results of the first correlational analysis (the group mean usage of each rhetorical feature of voice and the group mean essay scores from the TOEFL iBT independent writing tasks). Table 5.8 shows a positive correlation between the frequency of use of hedges and writing scores ( $r_s = .43, p = .34$ ) and a positive correlation between the frequency of use of direct reader references and writing scores ( $r_s = .39, p = .38$ ), and a negative correlation between the frequency of use of attitude markers and writing scores ( $r_s = -.64, p = .12$ ). None of the correlations in Table 5.8 proved to be statistically significant. Regarding RQ2, there are some correlations between the frequency of use of each individual rhetorical feature and the overall quality of writing with the intervention group. Appendix P illustrates the results of these analyses on scatterplots.

**Table 5.8**  
*Correlations, Means, and Standard Deviations of the Overall Quality of Writing and the Frequency of Use of Each Rhetorical Feature of Voice Across 49 Essays ( $n = 7$ )*

	Quality of writing			
	$r_s$	Sig. <sup>a</sup>	<i>M</i>	<i>SD</i>
Quality of writing	—	—	2.8	.34
Hedge	.43	.34	2.29	1.02
Booster	.07	.88	7.11	1.38
Attitude marker	-.64	.12	2.55	.71
Authorial self-mention	.07	.88	10.58	4.04
Direct reader reference	.39	.38	7.01	2.3

<sup>a</sup> Asymp. Sig. (2-tailed).

Table 5.9 shows the correlations between the group mean usage of each rhetorical feature of voice and the group mean voice scores collected from the TOEFL iBT independent writing tasks. As with the data presented in Table 5.8, none of the correlations in Table 5.9 proved to be statistically significant, presumably because of the small number of participants

and the large number of compositions evaluated. Nonetheless, Table 5.9 shows that there was a positive correlation between the frequency of use of boosters and overall voice scores ( $r_s = .64, p = .12$ ) and between the frequency of use of direct reader references and overall voice scores ( $r_s = .64, p = .12$ ). There was a positive correlation between the frequency of uses of authorial self-mention and overall voice scores ( $r_s = .57, p = .18$ ). There was a negative correlation between the frequency of uses of attitude markers and overall voice scores ( $r_s = -.54, p = .22$ ). Regarding RQ2, in sum, there were varying degrees of correlations between the frequency of use of certain rhetorical features and the overall quality of voice in writing with the intervention group. Appendix Q illustrates the results of these analyses on scatterplots.

**Table 5.9**  
*Correlations, Means, and Standard Deviations of the Overall Quality of Voice in Writing and the Frequency of Use of Each Rhetorical Features of Voice Across 49 Essays (n = 7)*

	Quality of voice		M	SD
	$r_s$	Sig. <sup>a</sup>		
Quality of voice	—	—	2.87	.34
Hedge	-.11	.82	2.29	1.02
Booster	.64	.12	7.11	1.38
Attitude marker	-.54	.22	2.55	.71
Authorial self-mention	.57	.18	10.58	4.04
Direct reader reference	.64	.12	7.01	2.3

<sup>a</sup> Asymp. Sig. (2-tailed).

Radar graphs below visually display how each participant in the intervention group applied the rhetorical features of voice across their written compositions. Each spoke on the radar graph represents a data collection point, starting from the Baseline and then going clockwise. Each point on each spoke represents the number of uses per writing sample of a

particular target discourse feature by an individual student. Figure 5.1 documents Maru's uses of the target voice features. Maru's uses of hedges and attitude markers remained relatively constant throughout the intervention period, while there was a pattern similar to his use of hedges and attitude markers with his use of boosters and direct reader references, including a decrease in their use on Day 10 (from having used nine boosters on Day 8, to three boosters on Day 10, and then 16 uses on the posttest, and from having used 15 direct reader references on Day 8, to two direct reader references on Day 10, and then 14 uses of direct reader reference on the posttest). His uses of authorial self-mention were sporadic, with Day 4 representing their greatest use (24 uses).

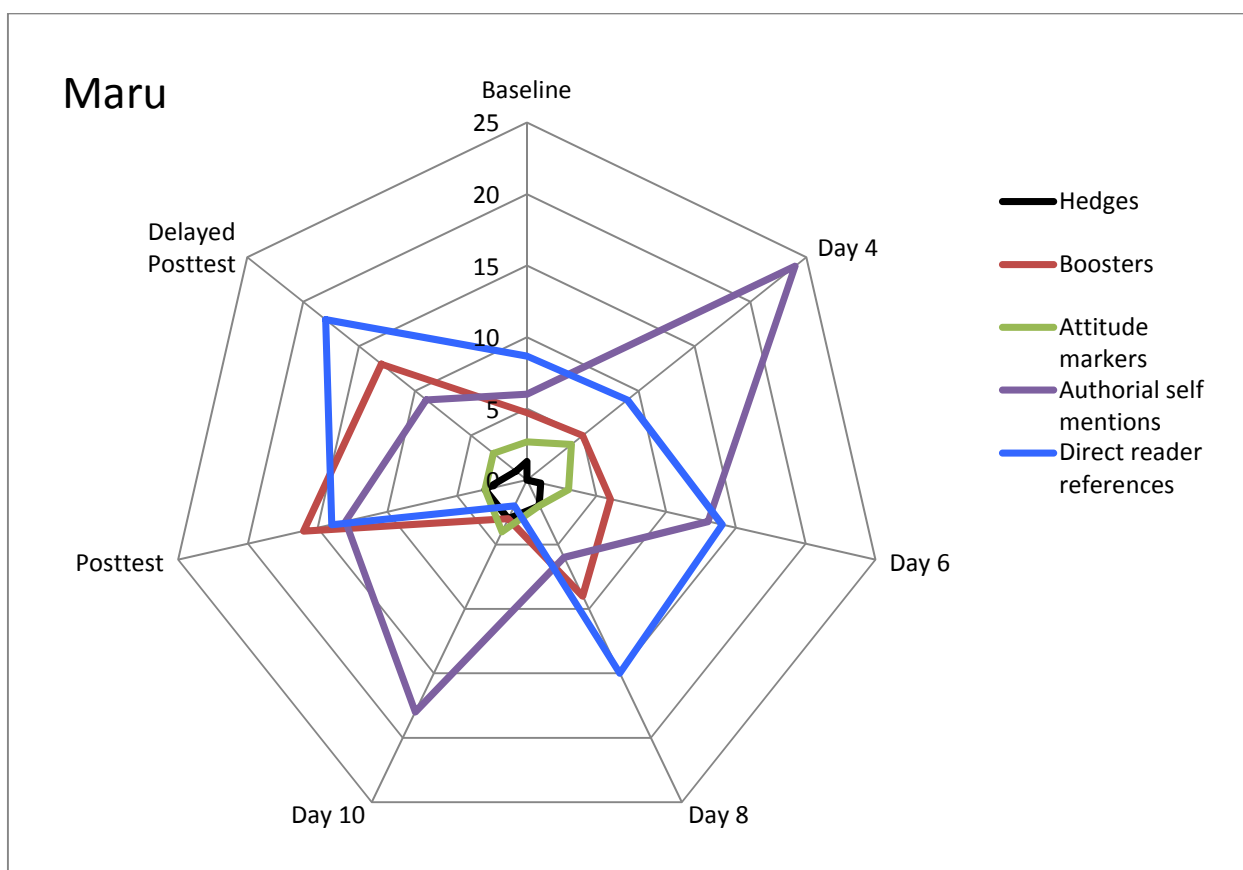


Figure 5.1. Radar graph depicting Maru's uses of each target rhetorical feature of voice.

Figure 5.2 documents Anne's uses of the target voice features. Notable is the posttest where there was an increase in the uses of boosters (from having used four boosters on Day 8

and Day 10, to ten uses on the posttest), authorial self-mention (from having used ten authorial self-mentions on Day 8 and on Day 10, to 23 uses on the posttest), and direct reader references (from having used nine direct reader references on Day 8 and four uses on Day 10, to 22 uses on the posttest), while Anne's uses of hedges declined over time (from five uses of hedges on Day 6 and on Day 8, to three uses on the delayed posttest).

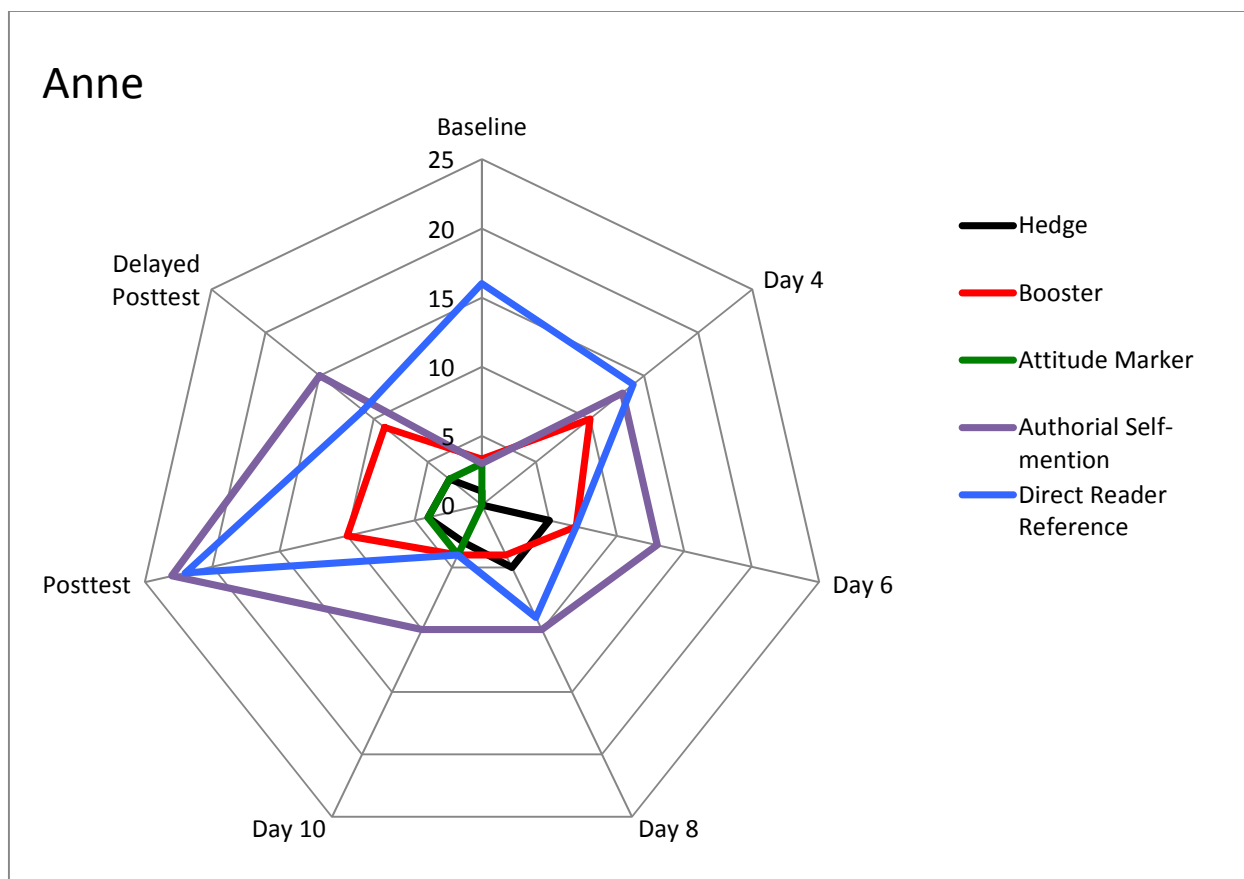


Figure 5.2. Radar graph depicting Anne's uses of each target rhetorical feature of voice.

Figure 5.3 describes Tom's uses of the target voice features. Tom's uses of hedges spiked on Day 10 (seven uses of hedges) but otherwise stayed relatively stable throughout the intervention period. Her uses of boosters (from 9.67 uses as her baseline average to six uses on the posttest), attitude markers (from 5.33 uses as her baseline average to one use on the posttest), and authorial self-mention (from 19.67 uses as her baseline average to 12 uses on the posttest) decreased over time, while her uses of direct reader references increased over



Days 4 (12 uses), 6 (ten uses), and 8 (14 uses) before declining close to her baseline level (3.3 uses) in the last three written compositions.

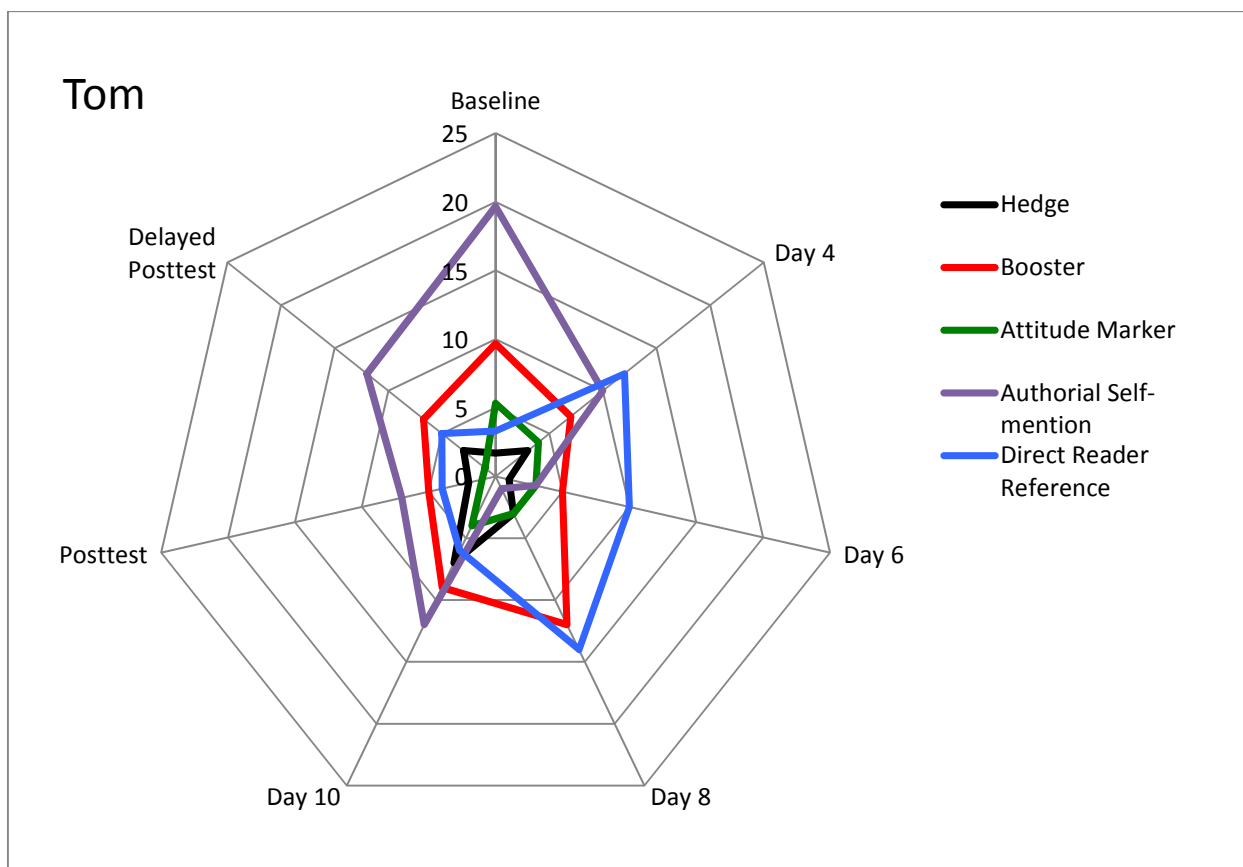


Figure 5.3. Radar graph depicting Tom's uses of each target rhetorical feature of voice.

Figure 5.4 documents Anny's uses of the target voice features. With the exception of Anny's consistent uses of boosters and near seldom uses of hedges, the remaining rhetorical features were employed sporadically with some extremely varied uses of authorial self-mention on Day 4 (27 uses) and Day 8 (one use).

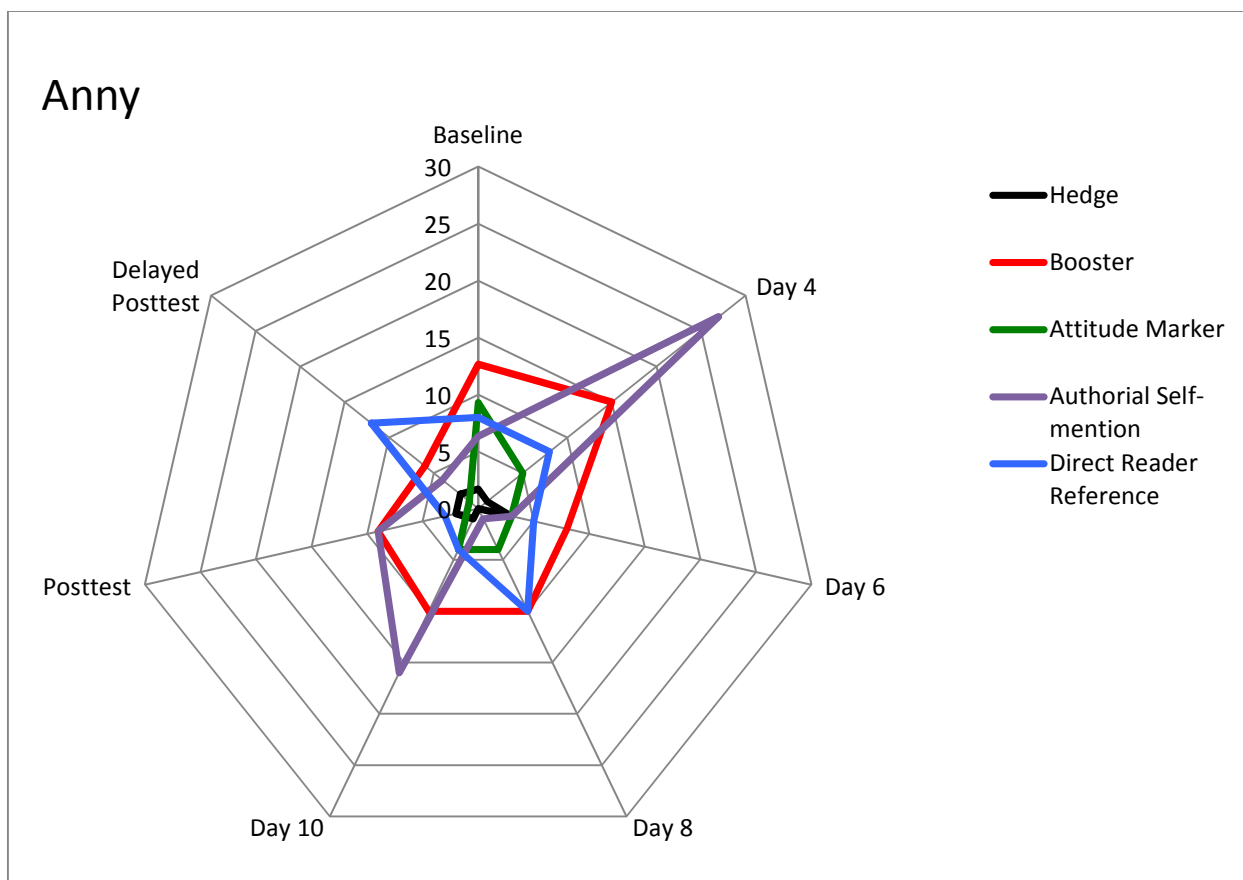


Figure 5.4. Radar graph depicting Anny's uses of each target rhetorical feature of voice.

Figure 5.5 describes Pi's uses of the target voice features. Excluding Pi's infrequent uses of attitude markers, her radar graph demonstrates an increase in uses of boosters (from a baseline average of 2.7 uses, to six uses in her delayed posttest) and uses of direct reader references (from a baseline average of 0.7 uses, to 13 uses in her delayed posttest) over time. Moreover, Pi's uses of authorial self-mention fluctuate, with extensive uses of this feature both at the start (18 uses on Day 4) and end (and 19 uses on the posttest) of the data collection phase of this study, while her uses of hedges are erratic (peaking on Day 10, with ten uses).

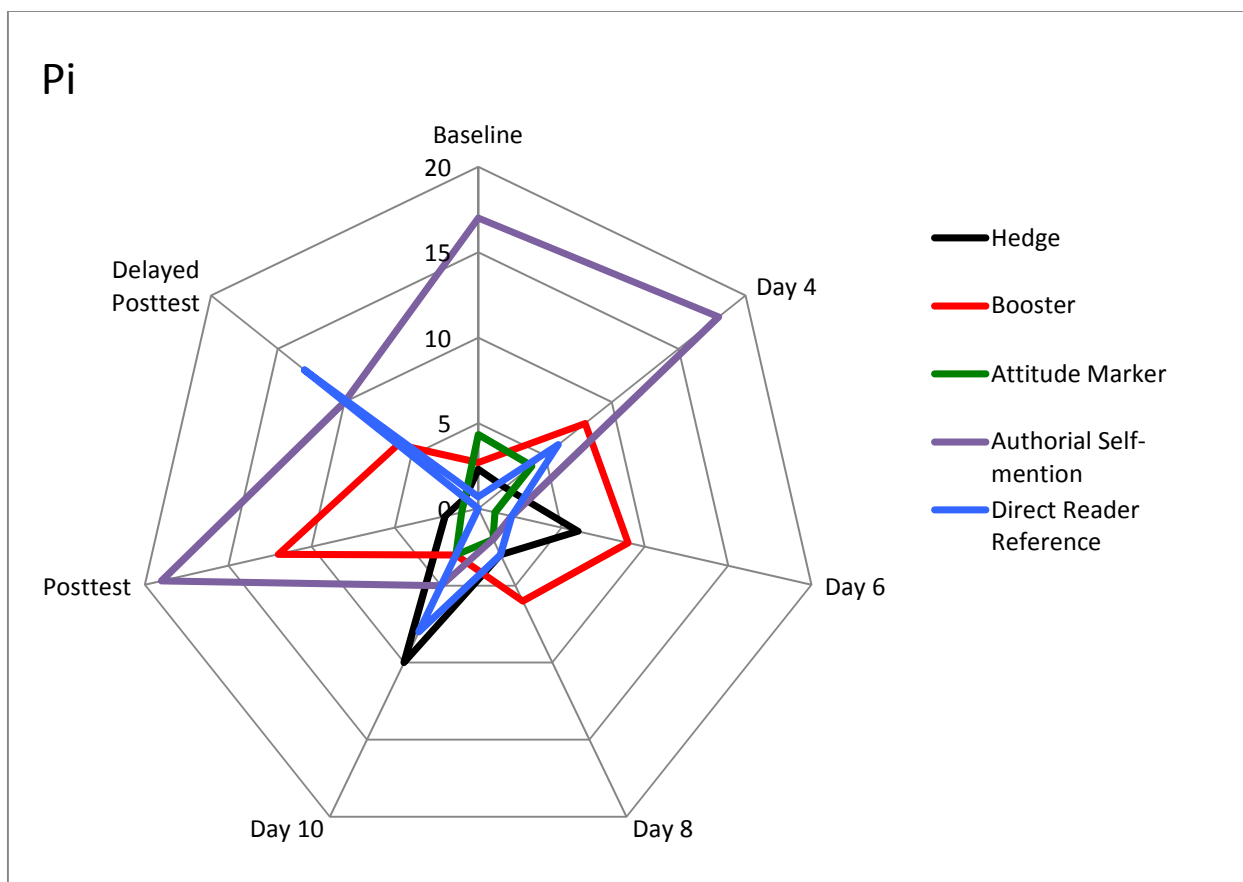


Figure 5.5. Radar graph depicting Pi's uses of each target rhetorical feature of voice.

Figure 5.6 documents Am's uses of the target voice features. Am's uses of hedges, attitude markers, and direct reader references were consistently low throughout the written compositions collected. Am's uses of boosters fluctuated during the study period (from a baseline average of 2.3 uses, to nine uses on Day 8, and two uses on the delayed posttest) while her uses of authorial self-mention showed some increase over time (from a baseline average of 9.33 uses, to 25 uses in her delayed posttest).

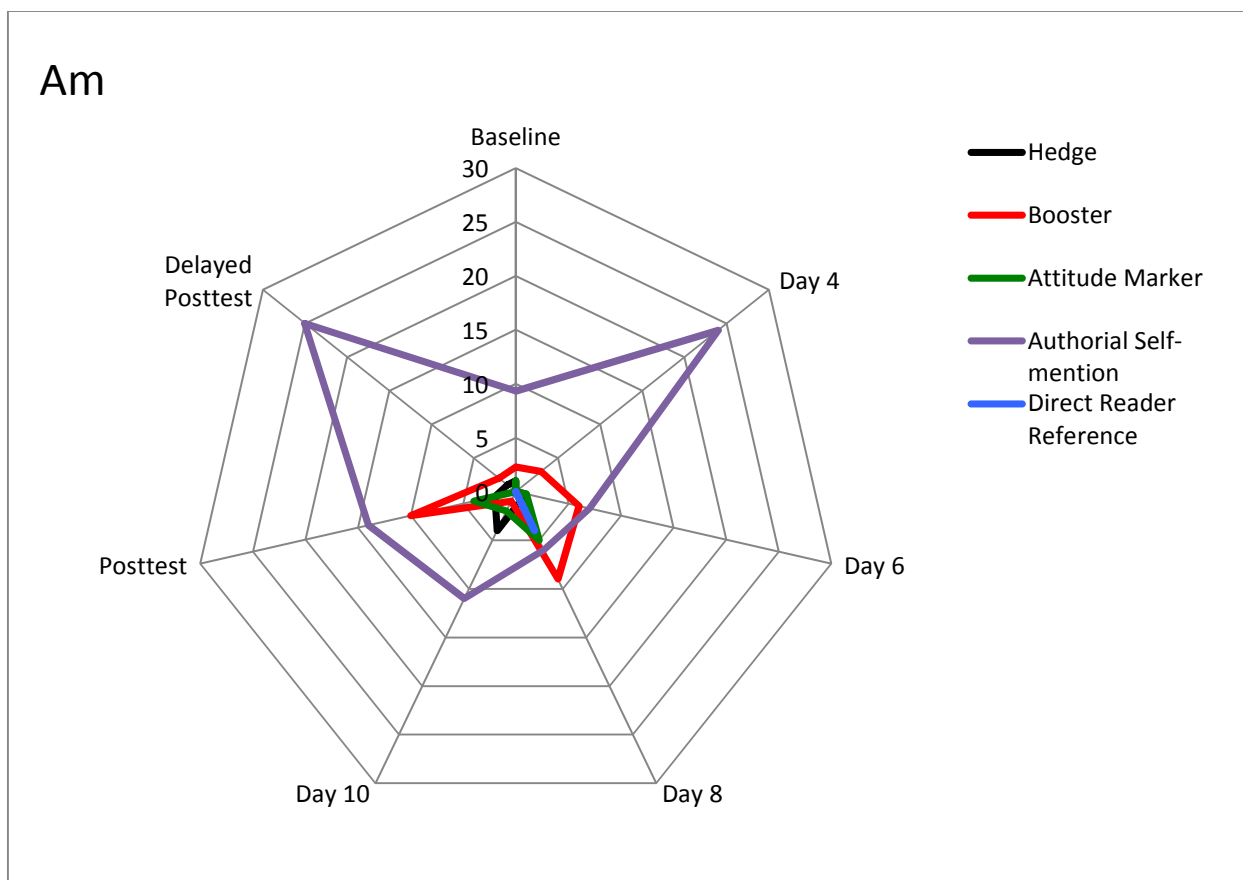


Figure 5.6. Radar graph depicting Am's uses of each target rhetorical feature of voice.

Figure 5.7 documents Angela's uses of the target voice features. Angela's uses of boosters and attitude markers were consistent throughout the time period (barring Day 10), while her use of authorial self-mention showed a decline during the last three written compositions collected (from a baseline average of 14 uses, then six uses on Day 4, and eight uses on Day 6, to one use on Day 10, two uses on the posttest, and four uses on the delayed posttest). Her uses of hedges were minimal, comparatively, and sporadic (four uses on Day 6 and no uses on Day 8); her uses of direct reader references increased (from a baseline average of 1 use, to 13 uses on the delayed posttest) but fluctuated throughout the time period.

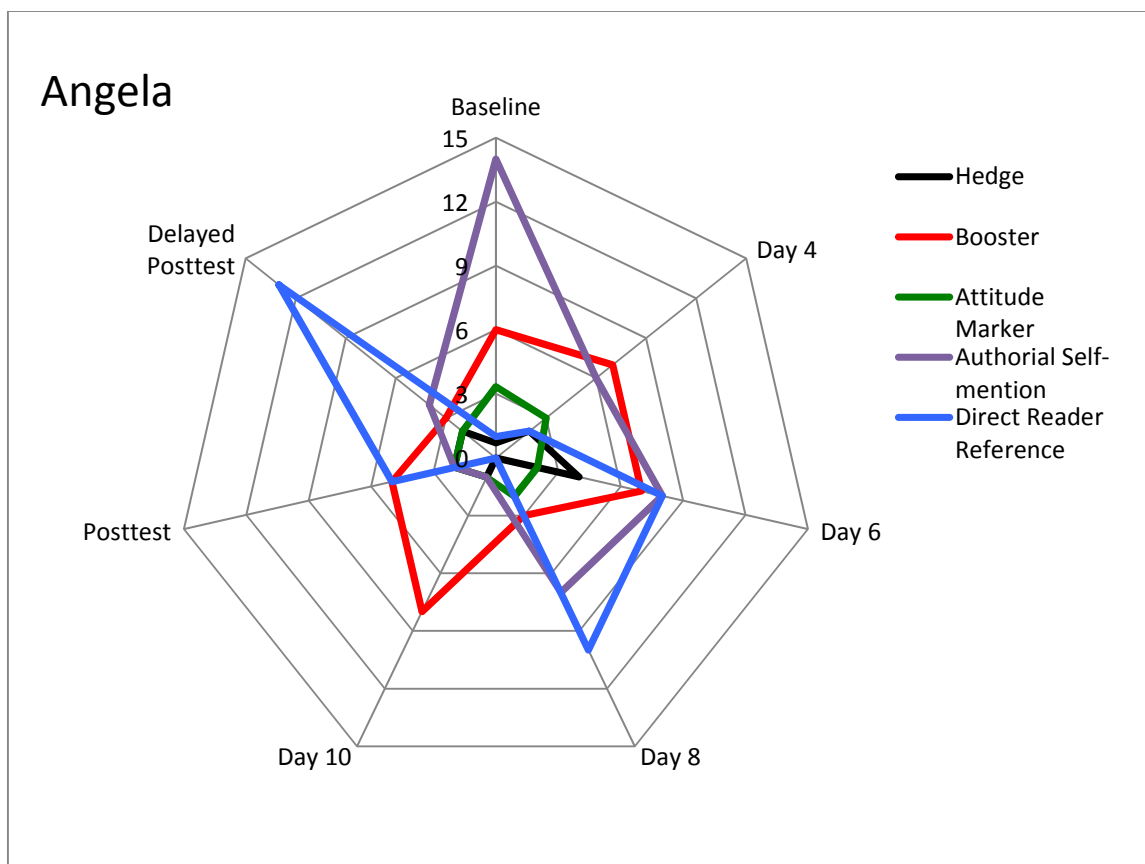


Figure 5.7. Radar graph depicting Angela's uses of each target rhetorical feature of voice.

These radar graphs indicate erratic uses of the rhetorical features across all of the participants. This consistently erratic behaviour emphasizes both a nonlinear path of cognitive uptake and a high degree of variability regarding L2 writing development across the participants. These factors contribute to understanding the path and variability of the microgenetic features of development demonstrated by this group of learners.

Figure 5.8 provides a graphic illustration of the mean uses of each rhetorical feature of voice across the participant group. This Figure demonstrates that authorial self-mentions, boosters, and direct reader references were used more frequently than hedges or attitude markers across most of the writing samples.

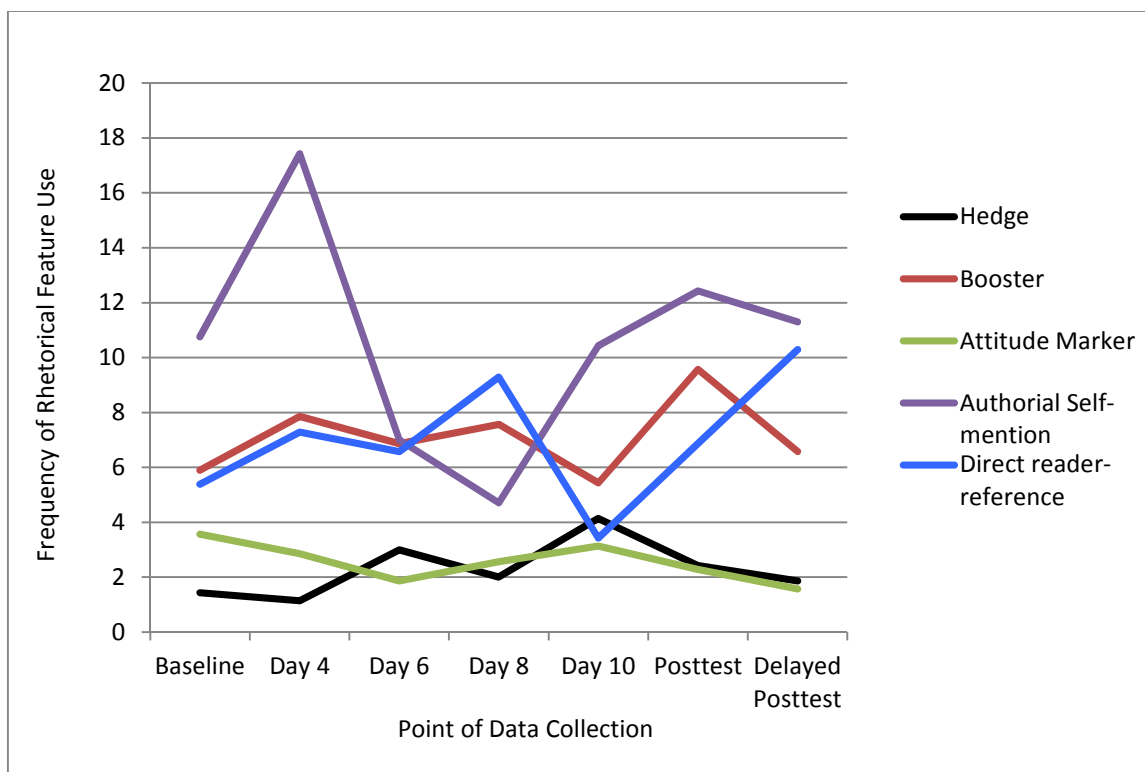


Figure 5.8. Group means depicting the use of each target rhetorical feature of voice across writing tasks ( $n = 7$ ).

### 5.2.1 Summary of findings for RQ2

Research Question 2 asked if there were direct relationships between each rhetorical feature of voice and the overall quality of writing and the overall quality of voice in writing across the writing samples collected for this study. The data suggest that there was not a generally consistent correlation between the number of rhetorical features used and overall writing scores and the type of rhetorical features used and overall writing scores. However, there were positive correlations between the number of certain rhetorical features employed (specifically, boosters, authorial self-mention, and direct reader references) by members of the intervention group in their writing and the voice scores on their compositions, although these correlations were not statistically significant.

**5.3 RQ3** *What is the relationship between (a) the quality of writing and (b) the quality of voice in writing?*

Spearman's rho was calculated to determine the relationship between group mean essay scores and group mean voice scores. The analysis indicated a positive correlation between the essay scores and the voice scores ( $r_s = .71$ ), although this correlation did not quite reach statistical significance ( $p = .07$ ). Appendix R illustrates the relationship on a scatterplot.

Verbal data related to this positive correlation appeared in learners' comments across four categories: (a) learners' conscious engagement with readers; (b) learners' self-awareness of their authorial voice; (c) the influence of stylistics; and (d) stylistics-based, small-group discussions.

First, participants showed conscious intent to engage a reader through their writing. Verbal data regarding conscious engagement were present across all the participants and across all of the target rhetorical features. For example, Tom discussed her use of an attitude marker: "I want reader knows about my opinion in short space so I put adjective to put more information about my opinion" (post-intervention interview, September 29, 2013). There was a particular focus on authorial self-mention (noted by three participants) and direct reader reference (noted by four participants). Comments showing a conscious effort by participants to engage readers included the following: Maru, discussing direct reader references: "I want the reader to think by himself or herself, so if the reader reads this sentence [*You might think that we can buy almost all products by online.*], they think [about] their own experience" (post-

intervention interview, September 29, 2013); Anne, discussing direct reader references: “At first time I think about topic but after three weeks I’m thinking about the reader [as I write] and I thinking I have to catch (+) I have to grab the reader’s mind so I concentrate on that [i.e., using direct reader references to engage the reader]” (post-intervention interview, September 29, 2013); Anny, discussing direct reader references: “To make my essay more strong because if I say ‘I’, ‘I’, ‘me’, ‘my’, just that kind of (+) my thing, it’s weak. It’s just my opinion but if I use ‘we’ or ‘our’ my essay will be more strong and more general for reader” (post-intervention interview, September 29, 2013); Tom, talking about authorial self-mention and direct reader references: “In most of my essay I talk about my experience so I use lots of ‘I’ and ‘my’, and also I wanted to mention about general information and other people’s ideas, including me, so [I] chose ‘we’ and ‘our’ too” (post-intervention interview, September 29, 2013).

Second, the participants also demonstrated self-awareness regarding the overall quality of voice manifest in their posttest writing samples. Specifically, this awareness supplements understandings of (a) participants’ knowledge of the function of the target rhetorical features, (b) how such awareness influenced participants’ employment of the target features to construct an authorial voice, and (c) how such awareness may have contributed to increased writing and voice scores (and thus, from a microgenetic perspective,



contributing to the source of change documented in learners' writing and authorial voice development). Table 5.10 summarizes participants' commentary in this regard.

Table 5.10

*Summary of Responses Regarding Participants' Interpretations of the Overall Quality of Voice in their Posttest Writing Task (n = 7)*

Interpretation of quality of voice	
	QI.B.3: Overall, but with respect to voice in particular, how would you describe the way you presented yourself in your writing?
Maru	Before I take this class I didn't use these (+) voice features, so this is completely different from before I wrote. [...] This essay is easy to understand because first is my experience and background. After that I write [with] hedge (+) and after that 'in my opinion' is easy to understand. 'Ah this is the writer's opinion' and 'absolutely' and 'definitely' means very strong opinion [that] I can understand. [...] I wanted to use more voice [features] [...] [This essay] could be better if I used more direct [reader] reference, I think.
Tom	I think the introduction and the conclusion I could use booster to say to mention about my opinion. But little bit I think in body paragraph I couldn't use booster or hedge or those [other voice features] very well so I think I should have written more example or more detail about me and my idea... the middle [of the essay] was really weak I think.
Anny	I think this essay's opinion is very strong because (+) this essay I can see 'really' or 'always' booster.
Angela	This essay sounds a little bit weak [author lacks confidence] so I should use more booster. (+) I think authorial self-mention is enough so if I use more another voice [feature] it can be better.
Pi	I just write my own story so I need to involve (+) involve to reader, like use 'you' or 'we'.
Am	Weak.
Anne	Voice (+) strong because I use many 'I' and examples and I tell about myself so strong voice.

Discounting Am, these comments demonstrated a conscious awareness of the function of the target rhetorical features while the essay samples provide data regarding the effectiveness of the target features in practice. As Section 5.1 also highlighted, every

participant attributed improvement in their overall writing to an increased awareness of how to employ the target rhetorical features of voice.

Third, comments by Maru, Tom, Anny, Pi, and Am presented already in Table 5.5 noted their senses of the influence of stylistics on developing their overall quality of writing and their overall quality of voice.

Finally, verbal data showed that group discussions influenced the concept development of the target rhetorical features and the overall quality of learners' writing. All seven participants remarked that discussing the target discourse markers influenced their understanding of (and use of) the five voice features. In addition, all seven participants observed that these discussions had a positive influence on the overall quality of their writing. The following samples underscore this influence:

- By discussing with other students they mention about these [voice features]. Maybe it imprint in my brain how it is good to use these voice [features] in my essay (Angela, post-intervention interview, September 29, 2013);
- To listen to other people's opinions is more helpful for me to write essay (Pi, post-intervention interview, September 29, 2013);
- It [the discussions] helped me to explain things [in my writing] in a more general way. [...] It's easy to come up the word (+) if you use it a lot when you speak (Anny, post-intervention interview, September 29, 2013);
- I think I was trying to use those voice features in discussion so it was easier to use those [in writing]

because in my mind I have some idea of these voice features. For example, hedge and booster. Sometimes I could write without thinking because I use [them] a lot in speaking from discussion or just in class talking (Tom, post-intervention interview, September 29, 2013); and

- Yea. (+) It was for sure effective (Maru, post-intervention interview, September 29, 2013).

The primary reason cited for how discussions contributed to developing authorial voice focused on the uptake of new information constructed from peers. This information included, for example, emphasis on target features that were missed or discussing alternative accounts of the function of a lexical item in question. Moreover, Maru remarked that the primary benefit of discussing the literary texts was motivation: the pressure he felt to discuss his findings motivated Maru to conduct thorough analyses of the texts. The end result, he noted, was having a better understanding of the target voice features.

### 5.3.1 Summary of findings for RQ3

A positive correlation was detected between learners' overall quality of writing and learners' overall quality of voice in writing, though the correlation did not reach statistical significance, presumably because of the small number of participants and compositions involved in this analysis. Verbal data from learners' comments across four categories (learners' conscious engagement with readers, learners' self-awareness of their authorial voice, the influence of stylistics, and stylistics-based, small-group discussions) highlighted

the dynamic and complex relation of influences that may help explain, in Chapter 6, this correlation.

**5.4 RQ4** *What is the relationship between (a) the frequency of use of each target rhetorical feature of voice and (b) the concept development of each rhetorical feature of voice?*

Spearman's rho was employed to answer RQ4. Table 5.11 displays the results of the analyses that examined the relationship between the use of each rhetorical feature and scores for the concept development of writing. (Concept development scores for writing and speaking tasks were derived from a coding framework adapted from Brooks et al., 2010 that scored concept development along a continuum, from spontaneous concepts to scientific concepts—see Table 1.1 for the concept development continuum, and see Subsections 3.4.2 and 3.5.4 for more on how concept development was operationalized.) Table 5.11 shows a positive correlation ( $r_s = .67, p = .22$ ) between the use of hedges and the concept development of hedges in writing, and a positive correlation ( $r_s = .7, p = .19$ ) between the use of direct reader references and the concept development of direct reader references in writing. There is also a negative correlation ( $r_s = -.4, p = .51$ ) between the use of authorial self-mention and the concept development of authorial self-mention in writing. (Appendix S illustrates these relationships on a scatterplot.) None of these correlations, however, proved to be statistically significant.

Table 5.12 displays the results of the analyses that examined the relationship between the uses of each rhetorical feature and the concept development of the rhetorical features in speaking tasks (Appendix T illustrates these relationships on a scatterplot). Table 5.12 shows a positive correlation ( $r_s = .62, p = .27$ ) between the uses of direct reader references and the

concept development of direct reader references in speaking tasks, and a negative correlation ( $r_s = -.5, p = .39$ ) between the uses of boosters and the concept development of boosters in speaking tasks. None of these correlations, however, proved to be statistically significant.

**Table 5.11**  
*Correlations, Means, and Standard Deviations of the Frequency of Use of Each Target Rhetorical Feature of Voice and the Concept Development (CD) of Each Target Rhetorical Feature of Voice Across 35 Writing Tasks (n = 7)*

Frequency of use of rhetorical feature	Concept development score for each rhetorical feature (writing)											
	1		2		3		4		5		M	SD
	$r_s$	Sig. <sup>a</sup>	$r_s$	Sig.	$r_s$	Sig.	$r_s$	Sig.	$r_s$	Sig.		
1. Hedge	.67	.22									2.29	1.02
2. Booster			-.05	.94							7.11	1.38
3. Attitude marker					-.15	.81					2.55	.71
4. Authorial self-mention							-.4	.51			10.58	4.04
5. Direct reader reference									.7	.19	7.01	2.3
M	3.09		3.83		3.11		4.11		3.57			
SD	.52		.07		.4		.22		.42			

<sup>a</sup> Asymp. Sig. (2-tailed).

See Figures 5.9, 5.10, 5.11, 5.12, and 5.13 for graphic representations of the relationships between uses of rhetorical features and the concept development of each rhetorical feature. (In viewing these Figures, note that the data collected during the intervention phase of this study commenced on Day 4, after the initial three days of the study served to introduce students to stylistic analyses.)

Table 5.12  
Correlations, Means, and Standard Deviations of the Frequency of Use of Each Target Rhetorical Feature of Voice and the Concept Development (CD) of Each Target Rhetorical Feature of Voice Across 5 Speaking Tasks ( $n = 7$ )

Frequency use of rhetorical feature	Concept development score of each rhetorical feature (speaking)											
	1		2		3		4		5		M	SD
	$r_s$	Sig. <sup>a</sup>	$r_s$	Sig.	$r_s$	Sig.	$r_s$	Sig.	$r_s$	Sig.		
1. Hedge	.26	.67									2.29	1.02
2. Booster			-.5	.39							7.11	1.38
3. Attitude marker					0	1					2.55	.71
4. Authorial self-mention							.15	.81			10.58	4.04
5. Direct reader reference									.62	.27	7.01	2.3
M	3.03		3.31		2.14		4.54		3.66			
SD	.34		.68		.37		.5		.57			

<sup>a</sup> Asymp. Sig. (2-tailed).

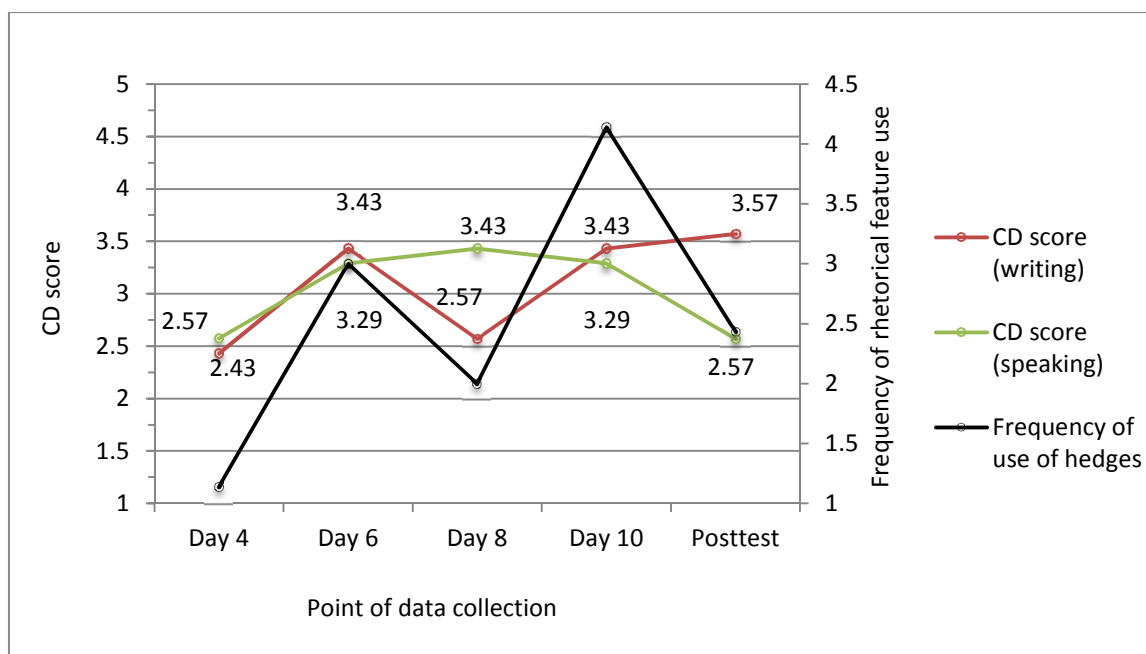


Figure 5.9. Relationship between mean concept development (CD) scores and frequency of use of hedges ( $n = 7$ ).

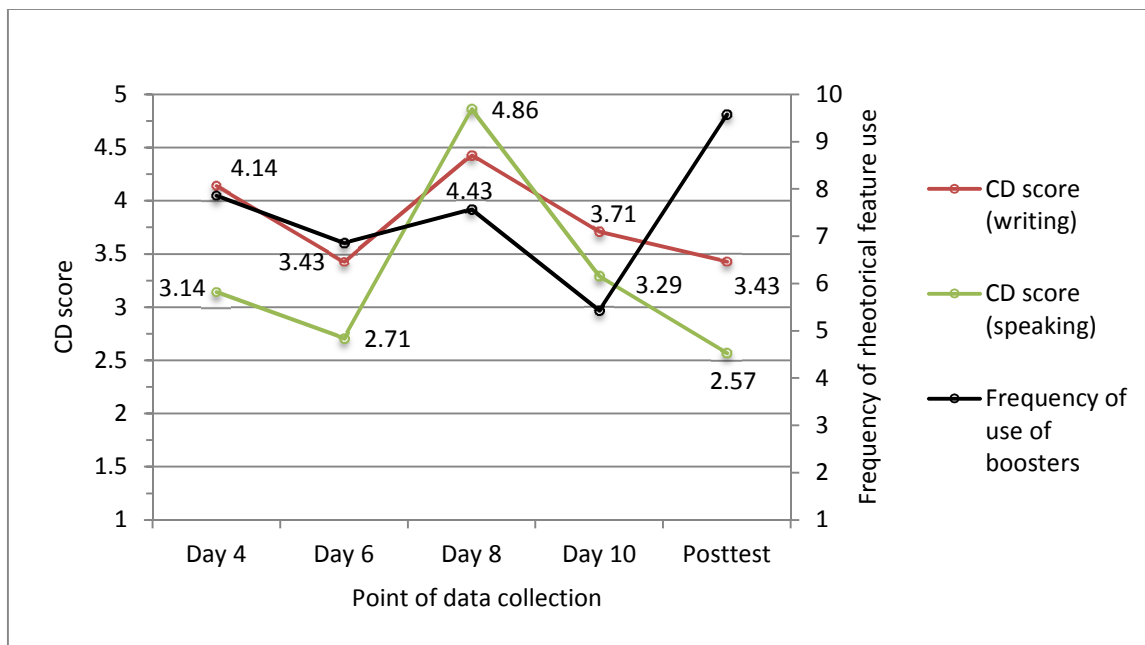


Figure 5.10. Relationship between mean concept development (CD) scores and frequency of use of boosters ( $n = 7$ ).

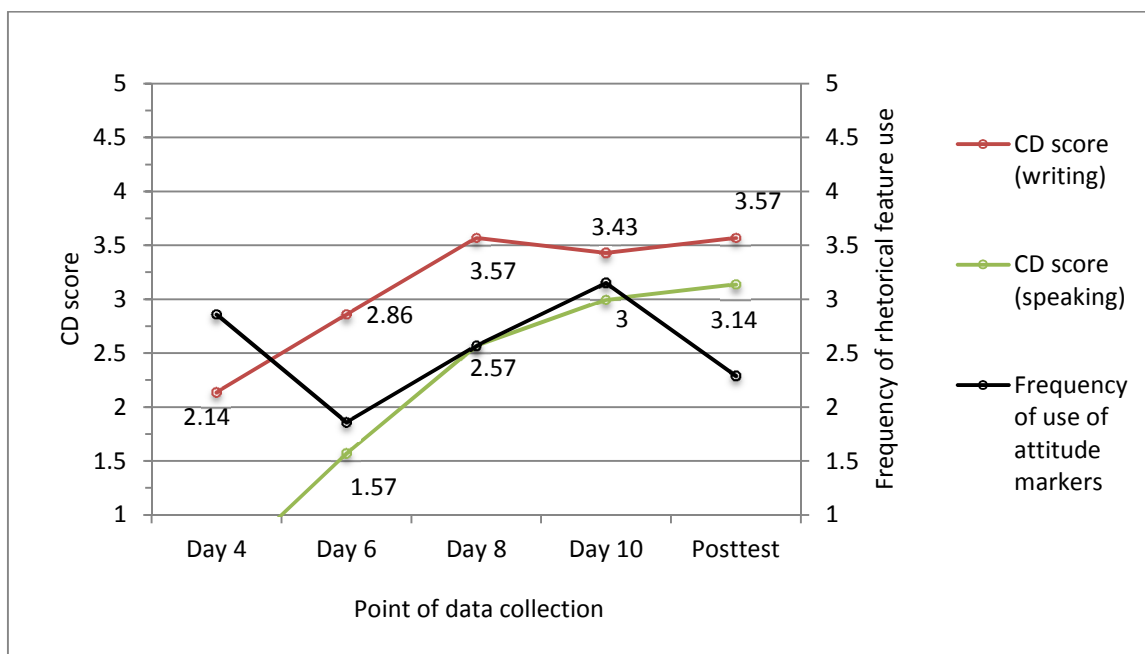


Figure 5.11. Relationship between mean concept development (CD) scores and frequency of use of attitude markers ( $n = 7$ ).

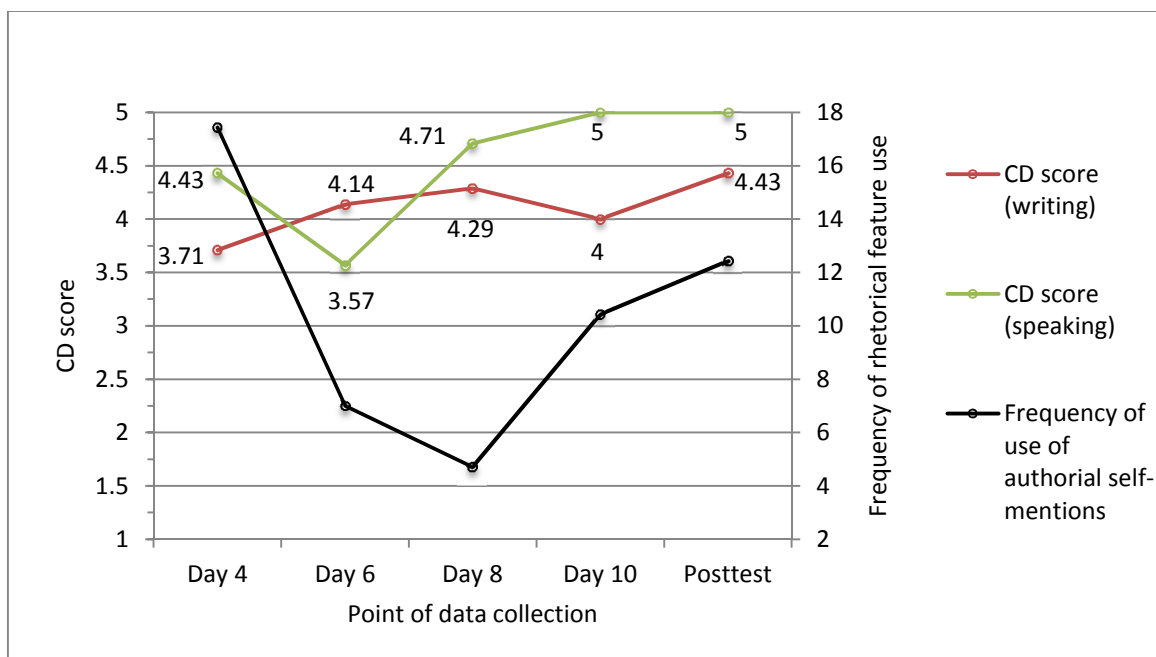


Figure 5.12. Relationship between mean concept development (CD) scores and frequency of use of authorial self-mention ( $n = 7$ ).

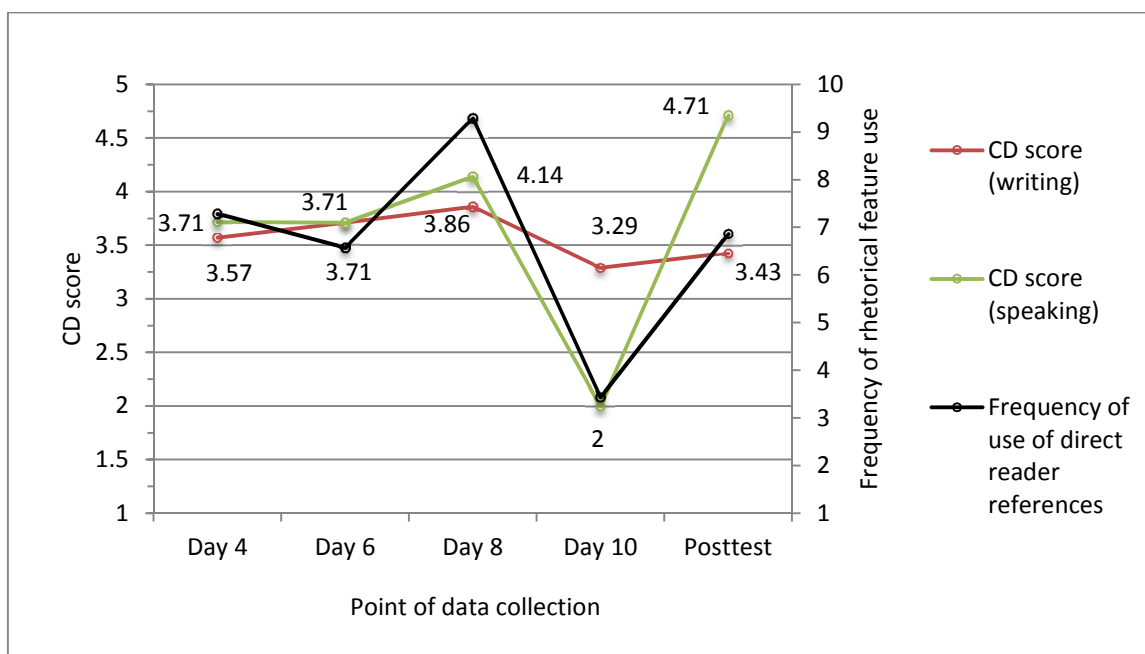


Figure 5.13. Relationship between mean concept development (CD) scores and frequency of use of direct reader references ( $n = 7$ ).



The verbal data documented in Section 5.3 highlight consistent commentary regarding learners' use of and awareness of the function of direct reader references. For example, as noted in Section 5.3, Anne expressed the following regarding her use and understanding of direct reader references: "At first time I think about topic but after three weeks I'm thinking about the reader [as I write] and I thinking I have to catch (+) I have to grab the reader's mind so I concentrate on that [i.e., using direct reader references to engage the reader]" (post-intervention interview, September 29, 2013). (See 5.3 for similar comments in this vein.) The comments provided by Anne, Maru, Anny, and Tom may provide insight into the positive correlation detected between the use of direct reader references (though less its frequency of use) and the concept development of direct reader references in writing and speaking tasks.

Regarding concept development across the target features, the statistical data are unable to provide a clear understanding of the relationship between the frequency of use of each rhetorical feature and the concept development of the target features. As the data suggest, there is a disconnect between participants' high concept development scores and their use of the target features. Angela, for example, demonstrated a clear understanding of authorial self-mention but, based on advice from her IELTS instructor, refrained from using this target feature in her last three written compositions. Similarly, Anny noted that occasionally she avoided using authorial self-mention at the risk of being repetitious.

The verbal data also provide insight into the cognitive processes informing concept (i.e., voice) development. These processes are highlighted below in verbal data from Tom

and Pi. As documented in Table 5.10, here Tom described her use of voice features in her posttest writing sample:

I think the introduction and the conclusion I could use booster to say, to mention about my opinion. But little bit I think in body paragraph I couldn't use booster or hedge or those [other voice features] very well so I think I should have written more example or more detail about me and my idea. (Tom, post-intervention interview, September 29, 2013)

The concept development of voice features, as Tom's comments and her written compositions indicate, followed an iterative path: In writing collected on Day 8 and Day 10 Tom was able to successfully employ the target discourse features; on Day 12 she was unable to, despite being cognizant of how such features could benefit her writing; in the delayed posttest (collected on Day 40) Tom was again able to successfully employ these features. Pi's written compositions and her comments regarding authorial voice development were similar: Pi noted that while writing she would often reflect on a literary text examined in class to help her select an appropriate rhetorical feature for her writing. Accordingly, Pi was retrieving information she had studied and was employing that information while she was writing. In both instances, cognitive processes displayed by Tom and Pi demonstrated an iterative pattern of cognitive uptake.

#### 5.4.1 Summary of findings for RQ4

Research Question 4 examined the relationship between the frequency of use of each target rhetorical feature and the concept development of these features across writing and speaking tasks. There were more rhetorical features that demonstrated positive correlations between rhetorical feature use and concept development in the writing tasks than in the speaking tasks, however none of these correlations were statistically significant. Verbal data suggest a disconnect between what the participants were able to conceptualize mentally and their uses of the target features in their written and spoken communications.

#### **5.5 RQ5** *Are there differences in (a) the quality of writing and (b) the quality of voice in writing between the control group and the course participants?*

To answer RQ5 I employed a Mann-Whitney U test to compare the quality of writing and the quality of voice in writing across the control group ( $n = 3$ )—whose tasks in this thesis study were to compose one baseline essay and to compose the delayed posttest essay—and the intervention group ( $n = 7$ ). The analyses compared written compositions that both groups composed: the first writing sample that informed the baseline data, and the last writing sample that constituted the delayed posttest. This configuration resulted in four different comparisons using Mann-Whitney U tests.

Table 5.13 displays the means, standard deviations, and medians for each set of scores that were analyzed. Table 5.14 documents the results for comparisons across the two groups regarding the quality of writing and the quality of voice in writing. As Table 5.14 shows, no statistically significant differences appeared across any of the four comparisons using the Mann-Whitney U tests. Regarding effect sizes, however, comparisons between the

quality of writing in the delayed posttest writing samples of both the control group and the intervention group produced moderate effect sizes ( $r = -.48$ ). Similarly, a moderate effect size appeared in comparisons between the quality of voice in the delayed posttest writing samples of both the control group and the intervention group ( $r = -.51$ ).

Table 5.13

*Means and Standard Deviations for Overall Quality of Writing Scores and Overall Quality of Voice in Writing Scores for the Control Group (n = 3) and the Intervention Group (n = 7) Across 20 Essays*

	Control group				Intervention group			
	First writing sample		Delayed posttest		First writing sample		Delayed posttest	
	Writing score	Voice score	Writing score	Voice score	Writing score	Voice score	Writing score	Voice score
<i>M</i>	2.43	2.3	2.6	2.43	2.23	2.21	3.17	3.11
<i>SD</i>	.55	.66	.53	.51	.47	.76	.39	.6
<i>Mdn</i>	2.7	2.2	2.8	2.3	2.1	2.3	3	3

Table 5.14

*Mann-Whitney Analyses for Differences in the Quality of Writing and the Quality of Voice in Writing Between the Control Group (n = 3) and the Intervention Group (n = 7) Across the First Writing Sample and the Delayed Posttest, Across 20 Essays*

Points of comparison between the control group and the intervention group	<i>U</i>	<i>z</i>	<i>r</i> <sup>a</sup>	Sig. <sup>b</sup>	<i>d</i> <sup>c</sup>
Quality of writing in first essay sample	8	-0.6	-.18	.58	-.37
Quality of writing in the delayed posttest	4	-1.5	-.48	.13	-1
Quality of voice in first writing sample	10.5	.0	1	0	0
Quality of voice in delayed posttest	3.5	-1.6	-.51	.11	-1.2

<sup>a</sup> Effect size ( $r = z/\sqrt{N}$ ). <sup>c</sup> Standardized effect size ( $d = 2r \div \sqrt{1 - r^2}$ )<sup>b</sup> Asymp. sig. (2-tailed).

### 5.5.1 Summary of findings for RQ5

There were no statistically significant differences either in the quality of writing or the quality of voice in writing between the control group and the intervention group across the first and last written compositions collected for this study, but moderate effect sizes did emerge from both of these analyses.

## 5.6 Findings related to microgenetic processes

This section of Chapter 5 consolidates the data outlined above that relate to the five components of microgenetic change. To exemplify the data informing these components, Table 5.15 documents data for the five dimensions of microgenesis in the learning behaviour of one participant, Tom.

The depiction in Table 5.15 of a microgenetic account of Tom's learning was typical of what occurred across all of the participants. That is, learning across the dataset was: erratic and nonlinear, initiated primarily through a stylistics-based pedagogy and by learners' awareness of the relationship between a strong authorial voice and overall writing scores, consistent (i.e., participants demonstrated nonlinear learning and participants suggested stylistics was the main factor influencing change), wide in its breadth (as all of the learners were able to transfer their learning across modes and mediums of expression), and gradual (and ongoing). The exception to one these descriptors of change is Am, whose development was unique in one limited area: her rate of learning. While the other six participants demonstrated similar, gradual rates of learning, Am's radar graph (Figure 5.6) suggested that her rate of learning may have been considerably more gradual than that of her peers. Am's

**Table 5.15**  
*Data Documenting Microgenetic Changes in Tom's Development*

	Description of microgenetic change
Path	The radar graph depicting Tom's uses of rhetorical features (Figure 5.3), her writing samples from Day 8, Day 10, the posttest, and the delayed posttest, and verbal data collected from Tom (Table 5.10) all indicate that Tom's path of development was erratic, iterative, and consequently nonlinear.
Source	Tom's verbal data (Table 5.5) emphasizes stylistic analyses of literary texts as the primary source of change for her writing and authorial voice development.  Supplementary verbal data from RQ3 (that indicates a strong, though not statistically significant correlation between writing scores and voice scores) suggest that Tom was aware that increasing her quality of authorial voice would contribute positively to increasing her overall writing score. Tom's heightened awareness here suggests that being cognizant of the relationship between writing and authorial voice is also a potential source of the changes noted in her writing samples.
Variability	The verbal data of Tom represented in Table 5.5 is consistent with her peers in noting that employing stylistics for improving authorial voice and for improving overall academic writing was both beneficial and the primary source of changes in their writing.  The radar graph depicting Tom's uses of rhetorical features (Figure 5.3), her writing samples from Day 8, Day 10, the posttest, and the delayed posttest, and verbal data collected from Tom (Table 5.10) all indicate that Tom was consistent with her peers in the developmental trajectory of her authorial voice development (i.e., development was nonlinear across all of the participants).  Data from RQ2 regarding correlations between the frequency of use of each target rhetorical feature and writing scores indicate that there was no correlation between the frequency of rhetorical features used or the type of features used and writing scores, and that the uses and development of the target discourse markers (as depicted in Tom's radar graph, Figure 5.3) were varied and erratic across learners, including with Tom (i.e., across the group, behaviour was consistently erratic)
Breadth	Based on Tom's posttest writing sample and her verbal data (specifically, recordings collected from the classroom-based discussions), Tom transferred her learning from one context (literary text analysis) to another context (high-stakes essay writing), and engaged the target features across modalities (reading, or analyzing, writing, and discussing). Tom's breadth of learning was wide.
Rate	Data from Table 5.6 suggest that Tom's overall quality of voice was not sustained over time. As Tom noted, the three-week intervention period was insufficient for completely developing her authorial voice, although her writing samples and verbal data suggest otherwise. Tom's rate of change was gradual (and ongoing).

radar graph noted that she employed, for example, direct reader references on only one occasion (Day 8) and failed to use individual target features across her written compositions on nine occasions across all of the writing samples. By comparison, and as depicted in the

radar graphs, Anne failed to use any one target feature on three occasions, Angela failed to use any one target feature on two occasions, Maru, Anny, and Pi failed to use any one target feature on one occasion, and Tom was able to use each feature at least once in every written composition. Similar results also featured across the speaking tasks: Am demonstrated less consistent and less frequent uses of the target discourse features during the in-class discussions than her classmates did. Am's behaviour, then, was unique in that the other participants, by comparison, were able to use the target features across the dataset more consistently (though still erratically).

## Chapter 6

### Discussion

In this chapter I first summarize the principal findings documented in Chapter 5 (see Appendix V for a table summarizing the findings), and then I discuss the significance of these findings for expanding understandings of authorial voice development, CBI, and pedagogical stylistics. This discussion is approached primarily through the lens of a sociocultural theory of mind interpretation of SLA, supplemented by a complex dynamic systems theory interpretation of SLA, and outlines the theoretical and pedagogical impact of this study's findings to all of these domains. I then note limitations of this study and I suggest avenues for continued research in this and similar research contexts. I conclude this chapter by remarking on the value of the main findings of this thesis study.

#### 6.1 Summary of findings for RQ1

The first research question asked: *Is there an increase, over the intervention period, in participants' (a) overall quality of writing, and (b) overall quality of voice in writing? Is any increase in the quality of writing and the quality of voice sustained after the intervention?* Statistically significant differences over the intervention period were detected regarding the overall quality of writing ( $\chi^2(6) = 20.86, p = .002$ , with a large effect size,  $W = .5$ ) and the overall quality of voice in writing ( $\chi^2(6) = 21.37, p = .002$ , with a large effect size,  $W = .51$ ). Post hoc analyses were only able to detect significant differences across pairwise comparisons concerning the overall quality of voice in writing. Calculations of



effect sizes across both sets of posthoc analyses (42 comparisons in total), however, revealed very large (in 17 instances) and large (in 6 stances) effect sizes.

These statistical analyses, coupled with the verbal data summarized below, provide insight into microgenetic components of change in learners' writing behaviour (as I noted in Subsection 1.2.4, microgenetic accounts of changes in learner behaviour are marked by five components: path, source, variability, breadth, and rate, as outlined by Calais, 2008, and Siegler, 1995, 1996). The statistical analyses and the verbal data concerning RQ1 provide insight into learners' breadth of learning. The improvements documented in the preceding paragraph highlighted statistically significant changes in learners' academic writing and quality of voice in academic writing over the short term. These changes suggest a wide breadth of learning: By employing the target discourse features to meet their writing goals, participants transferred learning from analyzing and discussing literary texts to composing academic prose. The verbal data also emphasize the role of learner awareness, with the exception of Angela and Anne, of the benefits of the transferability of the target discourse features from literary texts to their academic essay writing. This awareness may have served learners in expanding their breadth of learning.

The verbal data also provide insight into the potential source of the statistically significant changes in learners' writing behaviour. Table 5.5, summarizing learners' comments about the effectiveness of stylistics for mediating voice construction, emphasized stylistics as the primary source of the statistically significant changes documented in RQ1.

The last component of microgenetic change concerning RQ1 is variability. Participants all noted that stylistics was the key source of the changes to their writing; this

lack of variability across the seven participants adds to the claim that stylistics was the primary cause of the changes displayed by the participants.

Verbal data concerning RQ1 also addressed two specific characteristics of CBI: explicit instruction and the OBA. The participants suggested that explicit instruction in the form of teacher commentary contributed to—or was a source of—writing and voice development. Regarding the OBA, three learners commented positively on its effectiveness for developing authorial voice while four learners commented negatively on its role. Accordingly, the effectiveness of the OBA in this learning context is inconclusive, and the OBA cannot be relied on to help explain the statistically significant results documented in response to RQ1.

Thus far I have suggested that stylistics is the primary cause or source of the changes documented in RQ1. However, alongside verbal data on the effectiveness of sample essay analyses for modifying learners' writing (noted in Section 5.1), comments regarding explicit instruction and the OBA begin to reveal the complex dynamic that contributed to changes in learners' writing and authorial voice development. (In Subsection 6.6.1, I develop the idea of complex dynamic systems theory and I discuss how this perspective on SLA can help interpret some of the findings from RQ1 and, subsequently, the other research questions.)

Increases in the quality of participants' writing were sustained over the four weeks between the posttest and the delayed posttest, but increases in the quality of voice in writing were not sustained over this time. Given that participants were required, as part of the inclusion criteria for this study, to have had experience with standardized proficiency exams that include an argumentative writing task (e.g., the TOEFL or the IELTS), previous experience with the argumentative writing genre employed in this thesis study may have

helped participants retain changes in their writing scores. Moreover, these sustained changes may also be rooted in Japanese learners' understanding of deductive, argumentative rhetorical styles in their L1. For example, Japanese L1 rhetorical patterns overlap, to some degree, with English rhetorical patterns of written engagement. Kobayashi (1984) and Kubota (1992, 1997, 1998) both argue that deductive rhetorical patterns—of the type dominant in high quality responses to the TOEFL iBT independent writing task—are present in Japanese written prose (in about half of students' writing samples examined in Kubota, 1998). This tendency suggests that the participants may have been familiar with the rhetorical style of the argumentative essay. Accordingly, the intervention phase may have reinforced previous writing knowledge from participants' L1, or it may have made new knowledge of writing practices in the L2 easier to retain.

Regarding voice scores that were not sustained, a comment by Tom emphasized the complexity of authorial voice development, suggesting that the 16 instructional hours over 12 days of study may have been insufficient for internalizing the target voice features. Tom's performance in her written compositions, however, suggested otherwise, indicating a disconnect between her writing performance and her perceived proficiency. Moreover, and through the lens of microgenesis, development of authorial voice was gradual (and ongoing), rather than abrupt (the terms gradual and abrupt should be marked against an objective timeframe, but given the infancy of this research domain the present descriptors—borrowed from Calais, 2008—need to suffice until further studies are better able to quantify rates of change). While Tom suggested that an insufficient amount of time was the main reason she was unable to internalize the target features, other potential reasons include the quality of the instructional material (e.g., perhaps more time spent examining sample essays would have

proven more beneficial), the quality of the instructor (e.g., an instructor with more or varied experiences than me may have been able to effectively handle potential problems with learners' development during the intervention), or activities outside the classroom in which students may have or may have not engaged (e.g., during the four-week period between the posttest and the delayed posttest students may not have engaged in any writing practice). Moreover, studies have noted that L2 writing proficiency, at least at the level required to succeed in English medium, tertiary level studies, requires years to develop (Birrell, 2006; Leki, 2007; Sasaki, 2004); accordingly it is unsurprising that a three-week intervention period was not wholly sufficient for internalizing the target discourse features in this study. Last, the data suggest that the changes documented in participants' knowledge and writing appeared gradually, erratically, and variably rather than abruptly.

## 6.2 Summary of findings for RQ2

Research Question 2 asked: *What is the relationship between the frequency of use of each target rhetorical feature of voice, and (a) the overall quality of writing and (b) the overall quality of voice?* There was no clear pattern of correlation between the number of rhetorical features that participants used and the overall quality of their writing. This finding is consistent with earlier studies (Crossley & McNamara, 2012; McNamara, Crossley & McCarthy, 2010; Todd, Khongputb, & Darasawanga, 2007) that have documented the importance of the quality of language production over the quantity of language production. In contrast—and discounting the use of attitude markers—there were varying degrees of correlation (although none of these correlations were statistically significant) between the

overall quality of voice in writing and rhetorical feature use as well as variability from person to person and from time to time for writing for each person.

Figure 5.8 and the radar graphs (see Section 5.2) provide insight into voice development across two components of microgenetic cognitive processing: (a) regarding the path of cognitive processing, there seems to be a nonlinear path of concept development across all of the target discourse features (Section 5.4 expanded on this notion, and I provided further data therein that suggested that this path is iterative), and (b) regarding variability as a dimension of cognitive change, there is a lack of variability across the participants regarding the path of uptake (i.e., although each participant's uses of the target rhetorical features was erratic, this erratic behaviour was consistent across all of the participants).

Figure 5.8 also demonstrates the tendency for u-shaped behaviour (see Kellerman, 1985; Lightbown, 1983, 2000) in the use of authorial self-mention, the use of boosters, and the use of direct reader references. However, these tendencies are, perhaps, more likely explained by an effect curve rather than a learning curve, with the valleys and peaks in Figure 5.8 representing points of classroom intervention that focused on a particular rhetorical feature. For example, this is most certainly the case regarding the use of authorial self-mention. The classroom material leading up to the written composition collected on Day 4 focused on the use of this particular rhetorical feature.

### 6.3 Summary of findings for RQ3

The third research question asked: *What is the relationship between (a) the overall quality of writing and (b) the overall quality of voice in writing?* There was a positive

correlation between the overall quality of writing and the overall quality of voice in writing ( $r_s = .71$ ). But this correlation was not statistically significant ( $p = .07$ ), although the strength of the coefficient echoes Zhao's (2013) earlier findings that demonstrated a strong correlation between improved authorial voice and an increase in essay writing scores on the TOEFL test. Verbal data served to supplement these findings. First, learners observed that explicit instruction was effective for improving their senses of authorial voice. Second, while all of the learners noted the usefulness of stylistics for mediating authorial voice development—emphasizing from a microgenetic perspective, as in RQ1, stylistics as the source of the changes in learners' writing behaviour—five learners (Maru, Tom, Anny, Pi, and Am) also commented that they thought that stylistics was effective for developing their overall writing proficiency in English. Similar comments were made regarding the usefulness of the group discussions for improving both the concept development of the target voice features and the overall quality of learners' writing. Third, learners also observed that projecting a strong authorial voice could help improve their writing, and all learners made a clear effort to engage this element of their writing.

The awareness of reader engagement practices contributes to understanding the positive correlation between writing scores and voice scores. As a marker of concept development, improved awareness indicates an increase in knowledge regarding the function of the target features and knowledge of the conventions guiding their use. Verbal data suggest that the former (understanding the function of the target features) is the case across the participants, while statistical analyses of the writing samples suggest that the participants are capable of employing the features successfully. The link between an awareness of the function of each feature and the use of these features by the participants to improve their

authorial voice partially addresses why increases in voice scores correlate strongly with increases in writing scores. Furthermore, the presence of this awareness also suggests that, alongside stylistics, the extent of learners' awareness is a potential source for the changes documented in participants' written compositions: while the effects of learners' understanding of the form and function of the target voice features did not transpose fully into their writing, their enhanced awareness of the utility of the target features indicates some potential development or learning.

Attempts to explain the positive correlation observed in response to RQ3 have highlighted multiple factors that contribute to interpreting this finding. For example, the in-class discussions demonstrated how learners interacted with the target voice features, with their writing, and with each other to construct an authorial voice. These interactions, together with other influences observed in this section (learners' conscious engagement with readers, learners' self-awareness of their authorial voice, and the influence of stylistics), underscore the complexity of factors influencing this correlation and, subsequently, influencing the development of writing and authorial voice.

#### **6.4 Summary of findings for RQ4**

The fourth research question asked, *What is the relationship between (a) the frequency of use of each target rhetorical feature of voice and (b) the concept development of each rhetorical feature of voice?* More rhetorical features demonstrated a positive correlation between rhetorical feature use and concept development in writing tasks than in speaking tasks although none of these correlations were statistically significant. However, a lack of correlations overall indicates, again, that an increase in the number of rhetorical features that

learners employ does not equate with advancing concept development. Rather, as noted already, the quality of production (rather than the quantity of production) seems to be a more relevant marker of English writing development. Moreover, the high concept development scores that learners attained may indicate a disjuncture between what the participants were able to conceptualize and their actual use of the target features in their writing. That is, although learners thought they could use a particular feature effectively, they often avoided doing so in their writing, as with Angela and Anny. For example, in her last three written compositions Angela consciously avoided using authorial self-mention. Her decision reflected advice provided from an instructor unrelated to the present study. Anny also remarked that she sometimes avoided using authorial self-mention too frequently for fear of being repetitious. In short, the relationship between the frequency of use of the target voice features and their concept development is not clear.

Regarding microgenesis, learners demonstrated an iterative path of cognitive uptake as they engaged the target discourse features. This iterative process of development is consistent with a usage-based, emergentism theory of SLA (Eckstein et al., 2011; Hegelheimer, 2006; Hegelheimer & Fisher, 2006; Larsen-Freeman, 2011, 2013) that accounts for L2 development amid repeated and similar exposure to the target language feature(s). While providing insight into the microgenetic processes involved in authorial voice development (in particular, the path of cognitive uptake), these findings also help explain the erratic frequency of uses of the target rhetorical features and their relationship to concept development.



## 6.5 Summary of findings for RQ5

The final research question asked: *Are there differences in (a) the quality of writing and (b) the quality of voice in writing between the control group and the course participants?* There were no statistically significant differences between groups in the quality of writing and the quality of voice in writing, but comparisons between the quality of writing in the delayed posttest writing samples of the control group and the intervention group indicated moderate effect sizes ( $r = -.48$ ), and comparisons between the quality of voice in the delayed posttest writing samples of both the control group and the intervention group also produced moderate effect sizes ( $r = -.51$ ).

The data for RQ5 suggest that the intervention was not the sole cause of the statistically significant changes detected in learners' overall quality of writing in the long term and quality of authorial voice in the short term. However, this study did not aim to reveal causal connections—as the presence of a control group might otherwise suggest (Dörnyei, 2007)—but intended to explore a “unique perspective on the evaluation of intervention (or ‘treatment’) effects” (Glass, Willson, & Gottman, 2008). Accordingly, while the data for RQ5 may not support causal inferences regarding the effects of the intervention, collectively the findings from this thesis study, outlined below in Section 6.6, merit a larger, follow-up experimental study that would be more adept at revealing causal connections.

## 6.6 Theoretical and pedagogical implications

The findings from this research contribute new perspectives on authorial voice development, CBI-informed classroom practice, and pedagogical stylistics. First, by

expanding present conceptualizations of how learners develop their authorial voice, the findings help consolidate interpretations of authorial voice development through the lens of sociocultural theory; moreover, the findings also provide new insights into authorial voice construction through the lens of dynamic systems theory. Second, data from this study widen the context for CBI-informed teaching practices to include a high-stakes writing context that is informed by a stylistics pedagogy, and these data also raise important pedagogical questions about implementing CBI. Third, findings from this thesis study also support the value of stylistics-based interventions in this one L2 academic writing context, and these findings widen present conceptualizations of stylistics-informed learning.

#### 6.6.1 Implications for authorial voice development

As I noted in the introduction to this thesis study, one of the primary aims of this study was to document microgenetic changes in the voice development of Japanese learners as they navigated the complex task of constructing an authorial voice for high-stakes essay writing. To assist with the ensuing discussion, then, I begin this subsection by providing a brief review of the findings relevant to each dimension of microgenetic change. I then outline the implications of this work.

This study provides a glimpse into the path, source, variability, breadth, and rate (as outlined by Calais, 2008; Siegler, 1995, 1996) of authorial voice development exhibited by the participants in this study: (a) the path of cognitive uptake is nonlinear and erratic; (b) the source of the changes learners' exhibited—increases in the quality of authorial voice over the short term and increases in the overall quality of academic writing over the long term—derived from a composite of learning tasks and affordances that the participants engaged,

with particular emphasis on stylistic analyses of literary texts; (c) regarding variability, there was no variation regarding the causes of change in writing performance across the participants: all of the participants said that employing stylistics for improving authorial voice (and for improving overall academic writing) was both beneficial and the primary source of changes in their writing. Moreover, there was considerable variability regarding individual differences in patterns of cognitive uptake: all of the participants demonstrated erratic and hence nonlinear cognitive processing; (d) the breadth of learning is wide.

Learners transferred their learning from one context (literary text analysis) to another context (high-stakes essay writing), and engaged the target features across modalities (reading, or analyzing, writing, and discussing); and (e) the rate of change concerning concept development was gradual (see Table 5.15 for an example of how these features of change were manifest in one participant in this study). I discuss the implications of this microgenetic analysis of voice development through the lens of a sociocultural theory of mind interpretation of L2 writing development and SLA, and a complex dynamic systems theory of L2 writing development and SLA. I also discuss how findings related to uses of the target voice features corroborate with previous studies on voice in L2 writing, and the resulting implications. I expand on each of these below.

The non-linear path of cognitive uptake noted in this study makes three contributions to understanding L2 voice (and writing) development through the lens of a sociocultural theory of mind. First, this nonlinear behaviour corroborates earlier research that highlighted the nonlinearity and variability of developmental trajectories across learners of microgenetic learning events in L2 writing contexts (Cheng, 2010; Ferreira, 2005; Ferreira & Lantolf, 2008; Johnson, 2008; Park, 2012; Yanez-Prieto, 2010). Second, the findings reported here

expand understandings of the cognitive processes informing L2 writing development (that have typically focused on genre-based writing contexts, such as work by Ferreira, 2005, Ferreira & Lantolf, 2008, Johnson, 2008, and Yanez-Prieto, 2010) to include authorial voice development (and stylistics-informed learning contexts). Third, while a sociocultural theory of mind attempts to explain L2 development by conceptualizing sociocultural and cognitive factors that inform development, to date research into authorial voice construction has concentrated primarily on investigating only one component of this dyad: the socio-cultural—or what I designated as macrolevel—influences that inform development. Accordingly, this study contributes a microgenetic perspective on how L2 learners engage the process of voice development, thus adding a cognitive-oriented dimension to a sociocultural theory of mind interpretation of authorial voice development.

As discussed, the path of cognitive uptake described thus far contributes in three ways to a sociocultural theory of SLA; the path of cognitive processing thus described also has implications for a dynamic systems theory approach to SLA. As I noted in Section 2.4 of this thesis, and as data from this study demonstrate, sociocultural theory and dynamic systems theory may be interpreted as compatible and complementary theories of SLA. For example, the data on authorial voice construction inform the cognitive component of a sociocultural theory of mind interpretation of SLA by indicating the presence of nonlinear authorial voice construction whose developmental trajectory varies across learners. The same data underscore nonlinear development and varied developmental trajectories that characterize complex dynamic systems. In the paragraph that follows I expand this idea and then further develop the implications of this study for a dynamic systems theory of SLA.

In studies that examine dynamic systems there is widespread support for individual variability in L2 development (Baba & Nitta, 2014; de Bot et al., 2007, 2013; Larsen-Freeman, 2011; Spoelman & Verspoor, 2010; Verspoor et al., 2012; Verspoor, Lowie, & Dijk). The present thesis study substantiates this notion of individual variability as evidenced by the radar graphs in Section 5.2 and the verbal data in Table 5.5 (specifically within L2 writing development, as Baba & Nitta, 2014, Larsen-Freeman, 2006, and Verspoor et al., 2012, have demonstrated), and further expands the context for understanding variability within dynamic systems to include authorial voice development in a high-stakes writing context.

The data for this study also shed further light on the landscape of authorial voice development, with a focus on emergence and an emergentism theory of SLA. The statistically significant changes in learners' writing and quality of authorial voice during the intervention period, the erratic uses of the target voice features represented by the radar graphs, and the subsequent nonlinear path of cognitive uptake represent changes (and regarding the latter, uses) that are congruent with a usage-based, emergentism theory of SLA that emphasizes language development through repeated encounters with the target language feature(s) (see de Bot et al., 2007, 2013; Larsen-Freeman, 2006, 2011, 2013; Yetkin Özdemir, 2011), realized daily in this thesis study by the iteration of stylistic analyses. These factors (changes and uses) point to emergence, designated as patterns of “form–meaning–use composites” (Larsen-Freeman, 2011, p. 53)—present in participants' later writing samples—and corroborate similar findings regarding emergence in other learning contexts, such as those described in Baba and Nitta (2014), Eckstein et al. (2011), Hegelheimer (2006), Hegelheimer and Fisher (2006), and Larsen-Freeman (2012). While further research

(discussed in the next paragraph) would do well to corroborate the present findings regarding authorial voice development, from a pedagogical perspective, instructors may begin to consider the benefits of iterative learning contexts for constructing authorial voice.

However, I must also note that the design of this thesis study precludes certainty regarding the presence of emergence: In complex dynamic systems approaches to examining L2 development, the presence of emergence is defined by strict parameters (see Baba & Nitta, 2014) that this thesis study was not prepared to investigate. Accordingly, while this thesis study expands understandings of L2 writing development from an emergentism perspective of SLA to include authorial voice development, future studies are needed to corroborate some of the findings described in this study. In this vein, researchers can look to Baba and Nitta, Larsen-Freeman (2011), and Dörnyei (2014) for research methods (particularly, retrodictive qualitative modeling) that are adept at capturing emergence.

The microgenetic changes documented in this study also provide insight into the control parameter influencing emergence. If this parameter can be identified, “then we know what drives the system and are able to intervene” (Larsen-Freeman & Cameron, 2008, p. 54). As I have remarked above in this same subsection, the statistically significant increases in learners’ quality of authorial voice in the short term and overall quality of academic writing in the long term derived from a composite of learning tasks and affordances that included repeated opportunities to engage the target discourse features, across modalities, and with a particular emphasis on stylistic analyses within a CBI-informed pedagogy. This emphasis on stylistics, for which verbal data provides compelling evidence, suggests that stylistics was the primary control parameter in this learning landscape. This interpretation of the data has two implications. First, this interpretation offers support for the teaching of stylistics in similarly

placed learning contexts that focus on L2 writing development. Second, this interpretation also suggests (and corroborates previous research—Brooks et al., 2010, Ferreira, 2005, Ferreira & Lantolf, 2008, Negueruela, 2003, Swain et al., 2009, van Compernelle, 2011) that CBI is an effective pedagogy that can be employed to advance L2 proficiency and, within the context of this thesis study, to advance students' academic writing and authorial voice.

The data related to the microgenetic changes in learners' writing behaviour also indicate improvements in concept development of the five target rhetorical features. Table 5.15, which demonstrates Tom's development as exemplary of the group's general trajectory of learning, underscores these consistent changes (with the exception of Am). As I noted in Section 5.6 and based on my own classroom observations, Am's rate of learning appeared more gradual than that of the other participants. Given that emergence arises from repeated contact with the target language, Am's lack of uses of the target features—evidenced by her lack of use of individual target discourse features across her written compositions on nine occasions, and depicted in the radar graph (Figure 5.6)—may explain why her rate of change was slower than others. Potential explanations for Am's behaviour include that she was not yet familiar with how to employ the target voice features and thus required more time to understand how to employ them effectively (either because of her own learning trajectory or because of shortcomings with the teacher or the teaching materials), that she avoided using them for another reason that I was unable to discern, or that Am began this study at a lower English proficiency level (a TOEFL iBT score of 62) than other learners in this study had (the TOEFL iBT group mean score was 71.9)—although Pi commenced the study with the same proficiency level as Am, and Pi progressed throughout the intervention period with few similar difficulties due to factors that may be antithetical to those variables potentially

inhibiting Am's development—and thus received treatment that did not accord with her proficiency level. While scientific concept development of all the target features was not achieved across all of the participants, there was consistent development as evidenced by the microgenetic components of change noted in Table 5.15 and by improvements in writing scores over the long term and voice scores over the short term. The consistencies across the components of microgenetic change emphasize the processes (nonlinear, highly variable, stylistics-based, and gradual) informing change in L2 writing and authorial voice development through the lens of complex dynamic systems theory in one learning-to-write context.

Finally, this study makes one last noteworthy contribution regarding voice in L2 writing. Data for RQ3 revealed a positive correlation between the overall quality of writing and the overall quality of voice in writing ( $r_s = .71, p = .07$ ). This finding corroborates earlier work by Zhao (2013), who demonstrated that increases in authorial voice scores on the independent writing task of the TOEFL correlated positively with increases in essay scores. Given that Zhao's work was the first to note this correlation, showing this corroboration is important: By amassing studies that substantiate Zhao's findings, there is increased merit in instructors focusing on preparing students for the expectation of a clear authorial voice in the TOEFL independent writing task and perhaps other written composition tasks like it.

#### 6.6.2 Implications for concept-based instruction

My focus on CBI in this thesis study aimed to understand better CBI in an L2 writing acquisition context that is focused on voice construction. Specifically, below I discuss how this study initiates a conversation regarding the effectiveness of OBAs in this learning-to-



write context, how this study corroborates and then expands earlier findings that stress the import of languaging for concept development, and how this study advances CBI as a platform that engages a stylistics-informed pedagogy for advancing L2 writing and authorial voice development.

This thesis study initiates, and thus contributes, a discussion about learning outcomes with a potentially ineffective—as some participants in this study suggested—OBA. The OBA employed in this study (see Appendix L) was declared ineffective by four participants in this study, despite three of these participants achieving marked improvements regarding concept development across four of the five rhetorical voice features (see Section 5.1). Such improvements, given the perceived learner difficulties with the OBA, raise important questions about how students engage mediating artifacts in STI-informed learning contexts. These questions include the following: (a) In dynamic and multifaceted learning contexts, what exact contribution to learning does an OBA make? (b) Should instructors determine when an OBA is no longer beneficial for students and thus discard it, as Galperin (1969, 1992) suggested, or should learners decide when an OBA is no longer required, either because it proved effective (as with Maru, Tom, and Angela) or because its effectiveness is in doubt (as with Anny, Pi, Am, and Anne), thus constraining development? (c) Is an OBA a periphery mediational tool (as Anny, Pi, Am, and Anne suggested), or does an OBA directly facilitate learning (as Maru, Tom, and Angela noted)? (d) Are there are other mediating factors that learners are drawn to that may reconcile a perceived ineffectiveness with an OBA, thus allowing, as this study noted, Anny, Pi, Am, and Anne to improve as they did? These unanswered questions prompt concerns regarding the extent learners rely on an OBA to orient themselves to the target language. While answering these questions is beyond the

scope of this thesis study, one implication of stressing this issue is a call for studies that investigate answers to these questions. One potential avenue for pursuing these concerns is research through the lens of complex dynamic systems. Such investigations may account for other mediating factors (or other elements of a control parameter) that contribute to L2 learning, and such studies may account for variations in learners' reactions to different mediating tools.

Pursuing the questions outlined in the previous paragraph may also expand understandings of CBI-learning contexts in three other meaningful ways. First, further research may highlight how best to implement an OBA beyond our present controlled or stepwise understanding (Galperin, 1969, 1992; Johnson, 2008; Lantolf & Poehner, 2014). Second, continued studies may expand understandings of how L2 learners adapt to less than ideal learning conditions. This maneuvering and (re)interpreting of how best to apply STI in L2 settings is, moreover, consistent with CBI-informed praxis: sociocultural theory—and with it CBI/STI—is a “theory that simultaneously informs and responds to practice” (Lantolf & Poehner, 2014, p. 207). Accordingly, there is room to maneuver and to (re)design mediating tools to account for evolving and localized teaching and learning needs. Third, another contribution of this thesis study, then, calls attention to what sociocultural theorists have ignored or overlooked: how do instructors go about constructing effective OBAs that can facilitate L2 development? Lantolf and Poehner (2014) addressed this issue by calling on the descriptive (i.e., diagraming) practices of cognitive linguistics to support the designing of OBAs, but to my knowledge no studies or instructional material to date have provided L2 instructors with explicit directions on how to create an effective OBA (see, however—and out from under the purview of a sociocultural theory of mind approach to SLA—Liu, 2011

and Ojima, 2006 for suggestions on how instructors can guide learners to employ concept-mapping to facilitate L2 writing). Accordingly, this study contributes to this domain by noting this absence in published L2 studies, and by suggesting that the potential widespread success of STI-informed pedagogies in L2 learning contexts may benefit from research that examines this pedagogical concern.

This thesis study also makes two contributions to theories of languaging, a key phase of learning that mediates concept development (Brooks & Swain, 2009; Donato, 1994; Galperin, 1969, 1992; Lantolf & Poehner, 2014; Swain, 2006; Swain & Lapkin, 2002; Watanabe & Swain, 2007). In Section 5.3 I documented data that corroborate these theories, and these data also expand understandings of languaging to include its effectiveness for developing voice-related discourse features in high-stakes writing contexts. Second, this study expands theories of languaging to include acquiring metalanguage (see too Chang & Schleppegrell, 2011 and Vandrick, 2003) as a prerequisite to effective and efficient languaging that leads to concept development in first, stylistics learning contexts where CBI is practiced and second, potentially beyond. Present theories of languaging do not explicitly address the influence of metalanguage as a precursor to effective languaging. I argue that learners require a degree of comfort employing metalanguage—at least concerning stylistic analyses (e.g., while I did not document this in Chapter 5, ease using jargon such as *booster* or *stanza* during discussions later in the intervention period allowed for faster referencing opportunities, and hence, more efficient languaging opportunities).

Continued research may consider exploring the amount of time learners require to develop metalanguage and what activities might aid this process. From a pedagogical perspective, instructors may consider taking extra care when designing languaging

opportunities for L2 learners in stylistics-based learning contexts: Sufficient time should be allocated for acculturating students to the metalanguage required to language effectively.

Moreover, the idea of languaging makes salient the link between sociocultural theory and dynamic systems theory as both approaches to SLA interpret language development as emerging from individuals interacting. Accordingly, the concept of languaging described in the preceding two paragraphs also expands notions of interaction and emergence in dynamic systems theory to include languaging as a catalyst for and component of authorial voice development inside a dynamic and complex L2 learning environment.

Last, and continuing with the perspective of dynamics systems theory, this study demonstrated that CBI provides a learning space where emergence is both possible and observable, and that CBI provides a learning space where a stylistic pedagogy, acting as a potential control parameter, can thrive. These observations, then, have two implications. First, a CBI-informed pedagogy provides researchers interested in understanding SLA through dynamic systems theory with a compelling—and new—landscape for tracking L2 development. Future studies should take advantage of this opportunity, given that a dynamic systems theory approach to SLA investigates, among other components of L2 learning, the interaction of socio-culturally generated concepts (i.e., language concepts such as rhetorical features of writing) that are the cornerstone of a CBI pedagogy. Such studies may widen the context of understanding for how CBI can be utilized for advancing L2 proficiency beyond present conceptualizations. Second, while researchers expand this context, instructors presently interested in using stylistics to teach L2 learners (specifically, to teach academic writing and authorial voice development) can consider doing so via a CBI-informed pedagogy.

### 6.6.3 Implications for pedagogical stylistics

This thesis study impacts pedagogical stylistics across several key areas. First, this study addresses Hall's and others' concerns about a lack of empirical data to substantiate benefits ascribed in published studies to stylistics for mediating L2 learning. Second, by investigating authorial voice development in an ESL, high-stakes writing context, this study expands research foci in this domain, thus (a) widening understanding of stylistics as a pedagogical tool across different learning contexts, and (b) creating an avenue for future studies. Third, this thesis study augments preconceived benefits of stylistic analyses for L2 learners by including automatized knowledge of the target language features as a possible learning outcome. Fourth, this study provides a basis for research and data reporting methods in continued pedagogical stylistics research. In addition, and out from under the umbrella of pedagogical stylistics, this study helps substantiate earlier findings that address the effectiveness of employing literature (and different modes of exploring literary texts) as a tool for mediating L2 development. I discuss each of these points below.

As Hall (2005, 2007, 2014) and others have noted, research into pedagogical stylistics has yielded insufficient data to substantiate the widespread use of stylistics in L2 contexts. This thesis study contributes primarily to this issue: this study provides empirical data that help widen understandings of stylistics as a mediating tool in L2 classrooms and provides support for stylistics-informed interventions that facilitate authorial voice and academic writing development.

This study also expands the scope of stylistics research to date by focusing on academic writing production in a high-stakes writing context and by investigating ESL learners. Neither of these areas of study has been investigated extensively by researchers

interested in stylistics (Fogal, 2015; Hall, 2014). By expanding research foci, then, this present work generates a more complete picture of the effectiveness of stylistics across a varying degree of L2 learning contexts than was previously available.

The noted statistically significant changes in learners' writing and authorial voice development over the short term also augment Bellard-Thomson's (2010) taxonomy of learning outcomes for students employed in stylistic exercises. Bellard-Thomson's taxonomy included, for example, analyzing, synthesizing, and prioritizing techniques and the ability of students to "relate the data produced [from text analyses] to their understanding and explanation of the text(s)" (Bellard-Thomson, 2010, p. 55). This thesis study amends this list to include appropriation: Participants in this study demonstrated automatized competence of the target rhetorical features (in the vein of DeKeyser & Criado-Sánchez, 2013, or as I discussed in Subsection 6.6.1, a usage-based emergentism theory of SLA) triggered by repeated engagement with the target language. As I commented in Subsection 6.6.1, this iterative process proved partially effective for developing automatized competency (or for triggering emergence) of the target language features in this study. The presence of automatized knowledge, then, expands present conceptualizations of learning outcomes associated with stylistics beyond understanding and explaining—to empowering learners to manipulate language for their own expressive needs.

Out from under the umbrella of stylistics (and in closer proximity to literature-use theories that foreground the reader rather than the text—see Subsection 1.2.1), this study also enhances understandings of how studying literature (and different modes of exploring literary texts) can impact L2 learning (Belcher & Hirvela, 2000; Hanauer, 1997, 2001a, 2001b, 2011; Hirvela 2001, 2004, 2005; Iida, 2010, 2012, 2013; Kim, 2004; Kramsch, 2000;

Paran, 2008; Vandrick, 2003; Wang, 2009, among others) and, more specifically, L2 writing development (Belcher & Hirvela, 2000; Chen, 2006; Grabe, 2003; Hirvela, 2004, 2005; Iida, 2010; Vandrick, 2003). Enhancing the scope of this context is achieved by modifying Hanauer's (2011) recent theoretical position (built, arguably, on the work of Grabe, Hirvela, and Vandrick, as well as Hanauer's own extensive research in this area) concerning literacy transfer effects (viz., that writing skills can transfer from poetry writing to academic essay writing). This thesis study expands Hanauer's theory by having demonstrated that writing skills can transfer across genres via text analysis alone, without writing practice in the genre of the text under study. Moreover, while Hanauer's theory focuses on knowledge transfer originating in poetry, this study further expands Hanauer's theory to include knowledge transfer that originates in short stories as well.

This thesis study also expands the view of Grabe (2003), Hirvela (2004), and Vandrick (2003)—who have argued that the reading and writing of literature should be intertwined in L2 teaching contexts—by noting support for the concurrent teaching of reading (i.e., literary text analysis) and essay writing. This view is supported first by the proficiency gains made by learners in this study on their overall writing ability and their overall quality of authorial voice in writing, and second by verbal data. For example, Anny provided a particularly interesting comment in this regard:

But let's say songs and poems are songs and poems.  
They're not academic. ((Okay.)) No academic words. No  
however [transitions]. So [stylistics is] not negative  
but not helpful (+) I mean if I study just voice in  
stylistics I can't write a good essay so [I need]  
additional something. If I study academic base [as

opposed to only exploring literary texts] with the voice it's really nice.

(post-intervention interview, September 29, 2013)

Findings from this study, then, also address the legitimacy of pedagogical stylistics to complement academic writing courses.

This study also documented (see Table 5.10) that stylistics-based interventions succeeded in developing learners' analytic skills by engaging participants in conscious-raising activities. Learners were able to transfer analytic skills from literary text analysis to examining their own writing, thus expanding understandings of the role and importance of conscious-raising activities (see too Chang & Schleppegrell, 2011, Iida, 2010, Jwa, 2012, and Myskow & Gordon, 2012) to include stylistics-informed pedagogies.

Finally, as I remarked in Subsection 2.1.4, the research domain of L2 stylistics does not enjoy a history of research methods equipped to pursue future empirically-based studies. To this end, this thesis study contributes an analysis of the effects of a L2 stylistics-based pedagogy that is atypical in this research domain (see Iida, 2012, and Lin, 2010, for exceptions) and, accordingly, contributes to manners of collecting, analyzing, and reporting data that can serve future L2 stylistics-based research.

Based on the findings thus outlined in this subsection, instructors concerned with writing development and authorial voice construction in ESL academic writing contexts may consider the benefits of stylistics for mediating these ends. More specifically, instructors teaching in high-stakes learning-to-write programs like TOEFL test or IELTS test preparation courses may take advantage of the benefits stylistics bring to these contexts. Similarly, instructors may pursue using stylistics as a complement to (rather than focus on) classrooms aimed at teaching academic writing. For example, instructors may employ



stylistics alongside standard academic writing praxis, such as lessons on paragraph cohesion or essay structures, to enhance the L2 writing performance and the L2 voice development of learners. Instructors who employ stylistics or who may be interested in doing so may also consider the benefits of iterative learning environments for developing L2 writing proficiency and authorial voice. Finally, the transfer of analytic skills from examining literary texts to composing and examining academic prose may be helpful for instructors in, for example, genre-based writing programs where the importance of conscious-raising activities has long been stressed (Swales, 1990; Tardy, 2009) or in other learning-to-write contexts, as noted in Chang and Schleppegrell (2011), Iida (2010), Jwa (2012), and Myskow and Gordon (2012).

Thus far in this subsection I have discussed the impact of the findings from this thesis study on instructors and classroom instruction in stylistic contexts. Below I discuss the implications of these findings for researchers interested in this domain.

While this thesis study was able to provide some empirical data to help validate the usefulness of stylistics for L2 learning, the need for continued studies to address this issue remains. Accordingly, continued and extensive research that can expand the purview of L2 stylistics is the primary implication of this study for researchers interested in this learning context. For example, while this study investigated an ESL, high-stakes writing context, from the perspective of stylistic research, these two areas of study are novel, and future research may consider delving deeper into these (and other) study areas (see Fogal, 2015, for an extensive list, and Section 6.7, below, where I discuss avenues for future research) with the aim of presenting a better understanding of the effectiveness of stylistics across varied learning contexts than is presently known. To aid with continued studies, researchers may

draw from this study some examples of research practices that could benefit future investigations. Two such examples include having subjected data to statistical analyses and having employed mixed-method studies that used stimulated recall and semi-structured interviews to collect learner beliefs. Finally, and outside the domain of pedagogical stylistics, further studies aimed at investigating Hanauer's theory of literacy transfer effects may expand understandings of how literature can be implemented in L2 contexts.

### **6.7 Recommendations for future research**

This section of Chapter 6 directs potential future studies related to authorial voice development, CBI-informed L2 pedagogies, and pedagogical stylistics. In making these suggestions I also highlight some of the limitations of this study as a framework for further inquiry. Accordingly, I discuss both recommendations for future studies and limitations of this study in tandem.

This thesis study tracked cognitive processes informing authorial voice development across all five dimensions of microgenetic change (path, source, rate, breadth, and variability). In doing so, this study demonstrated that the cognitive demands of writing implicate multiple cognitive processes (e.g., planning, transcribing, and reviewing, as noted by Roca de Larios, Manchón, Murphy, & Marín, 2008; Ong, 2014; van den Bergh & Rijlaarsdam, 2001) across a complex and dynamic learning landscape. Moreover, this study provided in-depth analyses of two dimensions of change (the path and source of uptake) but failed to construct an in-depth analysis of the remaining dimensions (the rate, breadth, and variability of the changes observed). While this latter shortcoming may be overlooked (given that this study is (a) exploratory and (b) the first to examine voice development at the

microgenetic level across all five dimensions of cognitive uptake), future studies should consider providing a more focused account of voice development. For example, future studies may consider isolating particular dimensions of change for in-depth investigation (in the vein of P. Chang, 2010, and my earlier work, Fogal, submitted) or, alternatively, future studies may consider research designs that are adept at capturing all five dimensions concurrently. To this end, researchers may look to the advantages of investigating an intricate and multifaceted dimension of language development such as authorial voice construction through the lens of complex dynamic systems theory (e.g., see Dörnyei, 2014, or Larsen-Freeman, 2011, for suggestions on research design). Such research may help close the gap, for example, between macrolevel understandings of voice construction (reviewed in Subsection 2.2.1) and microlevel understandings of voice construction, with the aim of harnessing knowledge about authorial voice development for the benefit of L2 writing students.

This study also successfully employed stimulated recall interviews to capture the microgenetic processes highlighted above. However, attempts to reveal such processes as they unfold may be better served by concurrent think aloud protocols (van Someren, Barnard & Sandberg, 1994). Concurrent think aloud protocols address issues with post-hoc attempts to explain cognitive processing that may result in inaccurate participant assessment of learning behaviour (Gass & Mackey, 2000; van Someren et al., 1994). However, an aim of this thesis study was to assess change in the overall quality of voice in high-stakes writing contexts; accordingly, I decided that a concurrent think aloud protocol would have interfered with the writing task employed to measure change. Such interference would have raised concerns about construct validity (Janssen, van Waes, & van den Bergh, 1996; Smagorinsky,

1989). To offset potential inaccurate post hoc self-assessments of cognitive processes, the stimulated recall interviews were conducted within a limited period of time (three days after the writing task under question). Nonetheless, continued studies intending to measure microgenetic changes may wish to consider employing concurrent think aloud protocols alongside writing tasks that will not be adversely affected by this research instrument. Alternatively, future studies may consider implementing a series of staged interviews at intervals throughout an intervention phase (e.g., immediately after the writing tasks) to offset the potential of inaccurate post hoc self-assessments. Future studies can also employ screen recordings of learners engaged in writing tasks. Presenting these recordings during interview sessions may facilitate accurate post hoc self-assessments.

I successfully piloted the OBA that was employed during the intervention phase of this study, yet some participants still encountered problems conceiving of its benefits or encountered problems employing the OBA effectively, or both. In hindsight, further preparation—by expanding the pilot study and by further reviewing improvements to the OBA—may have generated an OBA that closely reflected Galperin’s intended use of STI. Learners would have also benefited from an avenue to express concerns about the OBA. This may have provided me, in the role of instructor, an opportunity to address issues or to make real-time changes to the OBA. Moreover, as noted in Subsection 6.6.2, there are no explicit materials available for L2 instructors that provide guidance in preparing an effective OBA. From the perspective of researchers, continued studies may consider exploring how best to design (and implement) an OBA as an effective mediating tool for L2 acquisition. To this end, while I was comfortable in the role of instructor, future studies may also consider employing teachers with different levels of experience implementing STI; such studies may

yield data on how best to implement and design different aspects of a STI-informed teaching and learning program. Instructors too may also profit from extended reviews of OBAs that they intend to use by seeking feedback from colleagues or examining relevant models of OBAs that presently exist. Time and resources permitting, instructors may also benefit from consulting their students about the content of the OBA (or, alternatively, students could design their own OBA with the support of an instructor).

By focusing on learners and the learner-text interface, the findings reported in this study also enhance theoretical understandings of how learners negotiate learning (manifest, in part, through authorial voice development) during classroom-based, stylistics-informed interventions. However, previous studies in pedagogical stylistics typically focused on the text under analysis (Hall, 2007) and the resultant interpretations, rather than on students and their interactions with the text. Further studies in L2 stylistics may wish to continue exploring this learner-text interface. Such investigations may help devise learning materials that harness the cognitive processes outlined herein (e.g., iterative learning activities in stylistic contexts may benefit from the templates or checklists promoted by Clark, 1996, McRae, 1997, and Simpson, 1997) or further studies may reveal undocumented learning processes that instructors can employ to facilitate L2 development. Moreover, continued studies may consider examining the outcome of stylistics-informed pedagogies in other language learning modalities. Collectively, then, continued studies may widen both the applicability of stylistics across varied L2 learning contexts and may help mold the teaching practices that inform such contexts.

Another strength of this study was that it reported on a single teaching and learning context. This focus, however, comes at the expense of understanding a wider educational

framework. Accordingly, future studies may consider expanding foci to include intervention periods of varying lengths, a less culturally and linguistically homogenous participant group and a larger sample size, learners at different stages of language development and age, as well as exploring different instances of authorial voice in different writing contexts.

Moreover, this study did not examine an intact classroom: Participants were assembled with the similar purpose of improving TOEFL writing scores and improving authorial voice.

Future studies, then, should also consider examining intact classrooms, and in L2 stylistic contexts can look to Plummer and Busse (2006), Warner (2012), and Zerkowitz (2012) as models. This study was also limited by its setting: a large, English L1 urban centre. The ubiquitous English language input participants enjoyed outside of the classroom may have influenced learner uptake and performance (Cumming, 2009, alluding to Carroll, 1977), and the present study was not designed to control or to account for these external variables.

Future research may consider similarly framed studies in foreign language contexts, or future studies may consider a research design that can account for such external variables.

This study was also able to highlight the importance of metalanguage for effective languaging. However, little is known about the degree of this importance (see Chang & Schleppegrell, 2011 and Vandrick, 2003), especially in the context of L2 voice construction. The field of SLA would benefit from studies that examine the extent of import ascribed to metalanguage and the cognitive and affective processes that learners engage as they acquire metalanguage as a tool for mediating L2 voice development.

Finally, this study also emphasized effect sizes in interpreting the dataset. Below I outline two specific considerations regarding this practice. The first concerns the role of

effect sizes in novel research contexts, while the second concerns employing the appropriate metric in L2 research for interpreting effect sizes. I expand these concerns below.

First, as Ioannidis (2008), Kline (2004), and Plonsky and Oswald (2014) discussed, studies in novel research areas (and when effect sizes are reported) are often prone to recording large effect sizes. This results from researcher biases that may manipulate variables within the study with the aim of producing large effect sizes that draw attention to, and merit, future studies. While there is an established tradition of research in L2 stylistics, findings drawn from statistical analyses in pedagogical stylistics are limited; in this regard, then, this thesis study presents a novel approach to analyzing data in this research domain and the concern thus documented needs to be addressed. Accordingly, in this study I marked a cautionary tone when large effect sizes were not accompanied by statistically significant findings. Moreover, while I took measures to dampen the effect of researcher bias (e.g., by recruiting experienced raters, by piloting the teaching materials, and by conducting a classroom-based study—as laboratory studies often produce larger effect sizes, as Plonsky and Oswald, 2014, noted), future, empirically-based studies in L2 stylistics should be aware of this tendency when designing research methods, and should caution against researcher bias.

Second, Plonsky and Oswald (2014) recently published an alternative metric for interpreting effect sizes in L2 research (see Subsection 3.5.1). Given the novelty of their metric (to date, no published studies in SLA have employed this scale), I decided that using this metric in this thesis study would be premature (although, as I suggested in Subsection 3.5.1, the scale that I did employ is similar to what Plonsky and Oswald recommended). Future studies, however, would do well to track the rise and usefulness of this metric for

interpreting effect sizes, thus allowing researchers to make an informed decision regarding its application.

### **6.8 Summary remarks**

This study focused on two L2 teaching and learning phenomena (pedagogical stylistics and a microgenetic account of voice development) while also widening the context of understanding and applicability of CBI. These three distinct research domains functioned in combination to provide insight into one learning-to-write context. Below I provide brief concluding remarks regarding each of these foci.

This thesis study makes the following contributions to understanding L2 voice development in a high-stakes writing context. From the perspective of Vygotsky's sociocultural theory of mind, this thesis study provides a microgenetic account of the cognitive processes informing authorial voice development across all five dimensions of microgenetic change. This microgenetic account of voice development contributes a comprehensive understanding of how learners engage the process of voice construction by adding another (microgenetic) dimension to an already robust collection of studies that have examined macrolevel influences on voice construction. This microgenetic analysis also emphasized the intricate, dynamic, and vast influences supporting the construction of authorial voice. Through the lens of complex dynamic systems theory, this thesis study expands understandings of learner variation in authorial voice developmental contexts and opens the door for continued studies into cognitive behaviour concerning L2 writing and authorial voice development. Through the same lens of complex dynamic systems theory, this thesis study also argued that stylistics can function as a control parameter in similarly



placed learning contexts, that CBI is effective in said contexts, and that iterative learning contexts implicate emergentist interpretations of voice construction that can support classroom interventions focused on authorial voice development. Last, this study corroborates initial work (Zhao, 2013) that demonstrated a positive correlation between increased voice scores and increased writing scores on the independent writing task of the TOEFL. The pedagogical implication of these findings include increased attention to developing iterative learning environments for developing authorial voice and academic writing, and increased attention to L2 writers' voice construction in high-stakes writing contexts.

This thesis study also contributes the following to CBI-informed theory and pedagogy. First, regarding theory informing CBI praxis, this study raises questions about the effectiveness of OBAs—either because they may be poorly-conceived or because learners do not perceive their (intended) effect—for concept development. I suggested that researchers investigate how they can provide instructors with tools for developing OBAs. Regarding languaging as put forth by Swain, this study corroborates present understandings of languaging and expands the context of its benefit for L2 learners to include languaging in stylistics-informed learning contexts that aim to improve authorial voice. This study also emphasizes the importance of acquiring metalanguage for languaging efficiently (and potentially effectively). I noted that continued research may consider exploring both the amount of time learners require to develop metalanguage and activities that promote metalanguage use in stylistic contexts. Instructors in similarly placed contexts should also consider taking care to ensure that learners have opportunities to develop relevant metalanguage. This study also demonstrated that CBI provides a new platform for examining

emergence from the perspective of a dynamic systems theory, which researchers should take advantage of. Last, I posited that stylistics-informed pedagogies can benefit from a CBI approach to teaching L2 learners, and I suggested that instructors interested in teaching with stylistics consider doing so via CBI.

Finally, this thesis study contributes the following to pedagogical stylistics theory and practice. This study provides empirical support for the effectiveness of stylistics for mediating authorial voice development in one L2 learning-to-write context, thus augmenting present expectations of learning outcomes derived from stylistics-informed studies to include appropriation of the target language. This study also contributes data collection, data analysis, and data reporting methods that are atypical in this research domain, and this study expands the scope of stylistics to include an ESL context and a high-stakes writing context focused on authorial voice construction. This study also modifies Hanauer's (2011) theory of literacy transfer effects in literature studies contexts by demonstrating the transfer of writing skills across genres via text analysis alone. Pedagogical considerations derived from this thesis study include heightened attention by instructors to the value of stylistics for improving authorial voice in high-stakes essay writing contexts and for developing analytic skills that may transfer into other (L2) learning contexts. Moreover, this study also highlights the value of literary text analysis as a complement to more traditional L2 academic writing courses. Last, instructors may also consider the benefits of iterative learning environments in stylistics-based pedagogies.

## References

- Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal*, 78(4), 465-483.
- Allen, H. W. (2009). A multiple literacies approach to the advanced French writing course. *The French Review*, 83(2), 368-385.
- Allison, D. (2004). Creativity, students' academic writing, and EAP: exploring comments on writing in an English language degree programme. *Journal of English for Academic Purposes*, 3, 191-209.
- Atwood, M. (1984). Doorway. Retrieved from: <https://www.tumblr.com/search/interlunar>
- Atwood, M. (1998). This is a photograph of me. Retrieved from:  
<http://www.poets.org/poetsorg/poem/photograph-me>
- Baba, K., & Nitta, R. (2011). Dynamic effects of repeating a timed writing task in two EFL university courses: Multi-element text analysis with Coh-Metrix. In P. M. McCarthy & C. Boonthum (Eds.), *Applied natural language processing and content analysis: Identification, investigation, and resolution* (pp. 397–411). Hershey, PA: IGI Global.
- Baba, K. & Nitta, R. (2014). Phrase transitions in development of writing fluency from a complex dynamic systems perspective. *Language Learning*, 64, 1-35.
- Bayyurt, J. (2010). Author positioning in academic writing. In S. Zyngier & V. Viana (Eds.), *Avaliações & perspectivas: Estudos empiricos em letras* (pp. 163-184). Rio de Janeiro: Publit.

- Belcher, D. & Hirvela, A. (2000). Literature and L2 composition: Revisiting the debate. *Journal of Second Language Writing*, 9, 21-39.
- Bellard-Thomson, C. (2010). How students learn stylistics: Constructing an empirical study. *Language and Literature*, 19(1), 35-57.
- Bellard-Thomson, C. (2011). Joining the stylistics discourse community: Corpus evidence for the learning processes involved in acquiring skills for stylistics analysis. In L. Jeffries & D. McIntyre (Eds.), *Teaching stylistics* (pp.51-68). London: Palgrave Macmillan.
- Bex, T. (1999, April). Towards a pedagogical stylistics. Paper presented at the Poetics and Linguistics Association Pedagogical Stylistics Special Interest Group, 19th Annual International Poetics and Linguistics Association Conference, University of Potchefstroom, South Africa.
- Biddix, J. P. (n.d.). Research rundowns: Effect sizes. Retrieved from <http://researchrundowns.wordpress.com/quantitative-methods/effect-size/>
- Birrell, B. (2006). Implications of low English standards among overseas students at Australian universities. *People and Place*, 14(4), 51-63
- Bloch, J. (2007). Abdullah's blogging: A generation 1.5 student enters the blogosphere. *Language Learning & Technology*, 11(2), 128-141.
- Bondi, M. 1999. *English across genres: Language variation in the discourse of economics*. Modena: Edizioni Il Fiorino.
- Bowden, D. (1999). *The mythology of voice*. Portsmouth, NH: Boynton/Cook.
- Bradford, R. (1997). *Stylistics*. London: Routledge.

- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.
- Brooks, L., & Swain, M. (2009). Linguaging in collaborative writing: Creation of and response to expertise. In A. Mackey & C. Polio (Eds.), *Multiple perspectives on interaction* (pp. 58-89). Mahwah, NJ: Erlbaum.
- Brooks, L., Swain, M., Lapkin, S. & Knouzi, I. (2010). Mediating between scientific and spontaneous concepts through languaging. *Language Awareness*, 19(2), 89-110.
- Bruner, J. (1983). *Child's talk: Learning to use language*. New York: Norton.
- Bukowski, C. (2002). trashcan lives. In *You get so alone at times that it just makes sense* (p. 23). Boston, MA: Black Sparrow.
- Bukowski, C. (2002). drive through hell. In *You get so alone at times that it just makes sense* (p. 122). Boston, MA: Black Sparrow.
- Byrnes, H. (2009). Emergent L2 German writing ability in a curricular context: A longitudinal study of grammatical metaphor. *Linguistics and Education*, 20(1), 50-66.
- Cabin, R. J., & Mitchell, R. J. (2000). To Bonferroni or not to Bonferroni: When and how are the questions. *Bulletin of the Ecological Society of America*, 81(3), 246-248.
- Calais, G. J. (2008). Microgenetic analysis of learning: Measuring change as it occurs. *National Forum of Applied Educational Research Journal*, 21(3), 1-7.
- Can, H. (2006). *An analysis of freshman year university students' argumentative essays*. (Unpublished Master's Thesis). Institute of Social Science, Boğaziçi University, Turkey.
- Carroll, J.B. (1977). The IEA studies in English and French as foreign languages: Implications for educational planning. *Studies in Educational Evaluation* 3, 1-5.

- Carter, R. (1997). *Investigating English discourse: Language, literacy and literature*. London: Longman.
- Carter, R. (2007). Foreword. In G. Watson & S. Zyngier (Eds.), *Literature and stylistics for language learners* (pp. vii-xi). Hampshire, UK: Palgrave Macmillan.
- Carter, R. (2010). Issues in pedagogical stylistics: A coda. *Language and Literature*, 19(1), 115-122.
- Carter, R. (2010). Methodologies for stylistic analysis: Practices and pedagogies. In D. McIntyre & B. Busse (Eds.), *Language and style: In honour of Mick Short* (pp. 55-68). New York: Palgrave Macmillan.
- Carter, R., & Nash, W. (1990). *Seeing through language: A guide to styles of English writing*. Oxford: Basil Blackwell.
- Chairlift (2008). Bruises. On *Does You Inspire You*. Brooklyn, NY: Kanine Records.
- Chang, P. (2010). *Taking an effective authorial stance in academic writing: Inductive learning for second language writers using a stance corpus* (Unpublished doctoral dissertation). The University of Michigan, Ann Arbor, MI.
- Chang, P., & Schleppegrell, M. (2011). Taking an effective authorial stance in academic writing: Making the linguistic resources explicit for L2 writers in the social sciences. *Journal of English for academic purposes*, 10(3), 140-151.
- Chang, Y. C. (2010). *Second language, literacy, and identity construction in online language learning* (Unpublished doctoral dissertation). University of California, Riverside, CA.
- Chen, Y. M. (2006). Using children's literature for reading and writing stories. *Asian EFL Journal*, 8(4), 210-232.

- Cheng, R. (2010). Computer-mediated scaffolding in L2 students' academic literacy development. *CALICO Journal*, 28(1), 74-98.
- Clark, U. (1996). *An introduction to stylistics*. Cheltenham, UK: Stanley Thornes.
- Clark, U. & Zyngier, S. (2003). Towards a pedagogical stylistics. *Language and Literature*, 12(4), 339-351.
- Coffin, C., & Donohue, J. P. (2012). Academic literacies and systemic functional linguistics: How do they relate? *Journal of English for Academic Purposes*, 11(1), 64-75.
- Coffin, C., & Hewings, A. (2004). IELTS as preparation for tertiary writing: Distinctive interpersonal and textual strategies. In L. Ravelli, & R. E. Ellis (Eds.), *Analyzing academic writing* (pp. 153-171). New York: Continuum.
- Coffin, C., & Hewings, A. (2005). Engaging electronically: using CMC to develop students' argumentation skills in higher education. *Language and Education*, 19(1), 32-49.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Erlbaum.
- Comparison of Effect Size Estimates (n.d.). *Proceedings of the Royal Society B: Biological Sciences*. Retrieved from <http://rspb.royalsocietypublishing.org/content/suppl/2009/03/20/275.1640.1305.DC1/rspb20080025supp01.doc>
- Connor, U. (2004). Intercultural rhetoric research: Beyond texts. *Journal of English for Academic Purposes*, 3(4), 291-304.
- Connor, U. (2008). Mapping multidimensional aspects of research. In U. Connor, E. Nagelhout, & W. Rozycki (Eds.), *Contrastive rhetoric: Reaching to intercultural rhetoric* (pp. 299-315). Amsterdam: John Benjamins.

Crain, J. C. (1993). Comment on “Two views on the use of literature in composition.”

*College English*, 55, 678-679.

Crossley, S. A., & McNamara, D. S. (2012). Predicting second language writing proficiency:

the roles of cohesion and linguistic sophistication. *Journal of Research in Reading*, 35(2), 115-135.

Cumming, A. (2001) Learning to write in a second language: Two decades of research. In

R.M. Manchón (Ed.), *Writing in the L2 classroom: Issues in research and pedagogy*.

Special issue of *International Journal of English Studies* 1, 1-23.

Cumming, A. (2009). The contribution of studies of foreign language writing to research,

theory, and policies. In R. Manchón (Ed.), *Writing in foreign language contexts:*

*Learning, teaching, and researching* (pp. 209-231). Clevedon, UK: Multilingual Matters.

Cumming, A. (2010). Theories, frameworks, and heuristics: Some reflections on inquiry and

second language writing. In T. Silva & P. K. Matsuda (Eds.), *Practicing theory in*

*second language writing* (pp. 19-47). West Lafayette, IN: Parlor.

Cumming, A. (2013a). Writing development in second language acquisition. In C. A.

Chapelle (Ed.), *The encyclopedia of applied linguistics*. Oxford, UK: Wiley-

Blackwell.

Cumming, A. (2013b). Multiple dimensions of academic language and literacy

development. *Language Learning*, 63(s1), 130-152.

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals,*

*and meta-analysis*. New York: Routledge.



- Davies D (1998) Metadiscourse and the evasive narrator: A process-based approach to teaching Ishiguro's *The Remains of the Day*. *Reading in a Foreign Language*, 12(1). 271-280.
- Davydov, V. V. (1988). Problems of developmental teaching: the experience of theoretical and experimental psychology research. *Soviet Education*. 30, 6-97.
- Davydov, V. V. (1990). Introduction: Lev Vygotsky and educational psychology. In L. S. Vygotsky, *Educational Psychology* (pp. xxi-xxxix). Boca Ratón, FL: St Lucie Press.
- De Angelis, G., & Jessner, U. (2012). Writing across languages in a bilingual context: A dynamic systems theory approach. In R. M. Manchón (Ed.), *L2 writing development: Multiple perspectives* (pp. 47-68). Berlin: de Gruyter.
- de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems approach to second language acquisition. *Bilingualism: Language and Cognition*, 10, 7-21.
- de Bot, K., Lowie, W., Thorne, S. L., & Verspoor, M. (2013). Dynamic systems theory as a theory of second language development. In M. Mayo, M. Gutierrez-Mangado, & M. Adrián (Eds.), *Contemporary approaches to second language acquisition* (pp. 199-220). Amsterdam: John Benjamins.
- DeKeyser, R. M., & Criado-Sánchez, R. (2013). Automatization, skill acquisition, and practice in second language acquisition. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Oxford, UK: Wiley-Blackwell.
- Donato, R. (1994). Collective scaffolding in second language learning. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 33-56). Norwood, NJ: Ablex.

- Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.
- Dörnyei, Z. (2014). Researching complex dynamic systems: 'Retrodictive qualitative modelling' in the language classroom. *Language Teaching*, 47(1), 80-91.
- Duff, P.A. (2008). *Case study research in applied linguistics*. New York: Erlbaum.
- Dylan, B. (1962). Oxford town. On *The Freewheelin' Bob Dylan*. New York, NY: Columbia Records.
- Eckstein, G., Chariton, J., & McCollum, R. M. (2011). Multi-draft composing: An iterative model for academic argument writing. *Journal of English for Academic Purposes*, 10(3), 162-172.
- Edmondson, W. (1997). The role of literature in foreign language learning and teaching: Some valid assumptions and invalid arguments. *AILA Review*, 12, 42-55.
- Educational Testing Service. (2005). *Setting the final cut scores*. Retrieved from [https://www.ets.org/Media/Tests/TOEFL/pdf/setting\\_final\\_scores.pdf](https://www.ets.org/Media/Tests/TOEFL/pdf/setting_final_scores.pdf)
- Educational Testing Service. (2008). *TOEFL® iBT Test independent writing rubric (scoring standards)*. Retrieved from [http://www.ets.org/Media/Tests/TOEFL/pdf/Writing\\_Rubrics.pdf](http://www.ets.org/Media/Tests/TOEFL/pdf/Writing_Rubrics.pdf)
- Educational Testing Service. (2011). *Supplementary comparison tables for TOEFL iBT® section scores*. Retrieved from [http://www.ets.org/s/toefl/pdf/supplementary\\_comparison\\_tables.pdf](http://www.ets.org/s/toefl/pdf/supplementary_comparison_tables.pdf)
- Educational Testing Service. (n.d.). *Writing topics*. Retrieved from <http://www.ets.org/Media/Tests/TOEFL/pdf/989563wt.pdf>
- Elbow, P. (1968). A method for teaching writing. *College English*, 30(2), 115-125.

- Elbow, P. (1994). What do we mean when we talk about voice in texts? In K. B. Yancey (Ed.), *Voices on voice: Perspectives, definitions, inquiry* (pp. 1-35). Urbana, IL: National Council of Teachers of English.
- Elbow, P. (1999). Individualism and the teaching of writing: Response to Vai Ramanathan and Dwight Atkinson. *Journal of Second Language Writing*, 8(3), 327-338.
- Ellis, N. C. (2011). The emergence of language as a complex adaptive system. In J. Simpson (Ed.), *Handbook of applied linguistics* (pp. 666-679). New York: Routledge.
- Ellis, P. D. (2010). *The essential guide to effect sizes*. Cambridge: Cambridge University Press.
- Erico, D. (2009). Me and you. Retrieved from <http://www.davidericowrites.wordpress.com/stories/>
- Ferreira, M. M. (2005). *A concept-based approach to writing instruction: From the abstract concept to the concrete performance* (Unpublished doctoral dissertation). The Pennsylvania State University, University Park, PA.
- Ferreira, M., & Lantolf, J. P. (2008). A concept-based approach to teaching writing through genre analysis. In J. P. Lantolf & M. E. Poehner (Eds.), *Sociocultural theory and the teaching of second languages* (pp. 285-320). London: Equinox.
- Fialho, O., Miall, & Zyngier, S. (2012). Experiencing or interpreting literature: wording instructions. In S. Csabi, L. Week, & J. Zerkowitz (Eds.), *Pedagogical stylistics: Current trends in language, literature and ELT* (pp. 58-74). London: Continuum.
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage.
- Fischer, K. W. & Granott, N. (1995). Beyond one-dimensional change: Parallel, concurrent,

- socially distributed processes in learning and development. *Human Development*, 38, 302-14.
- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76, 378-383
- Fløttum, K. (2012). Variation of stance and voice across cultures. In K. Hyland & C. S. Guinda (Eds.), *Stance and voice in written academic genres* (pp. 218-231). New York: Palgrave Macmillan.
- Fogal, G. G. (2015). Pedagogical stylistics in multiple foreign language and second language contexts: A synthesis of empirical research. *Language and Literature*, 24(1), 54-72.
- Fogal, G. G. (submitted). *A microgenetic analysis of authorial voice development in L2 academic writing*. Manuscript submitted for publication.
- Fogal, G. G. (2010). EFL literature studies: Student feedback on teaching methodology. *Asian EFL Journal*, 12(4), 114-129.
- Gajdusek, L., & van Dommelen, D. (1993). Literature and critical thinking in the composition classroom. In J. Carson & I. Leki (Eds.), *Reading in the Composition Classroom: Second Language Perspectives* (pp. 197-215). Boston, MA: Heinle & Heinle.
- Galperin, P. I. (1969). Stages in the development of mental acts. In M. Cole & I. Maltzman (Eds.), *A handbook of contemporary soviet psychology* (pp. 248-273). New York: Basic Books.
- Galperin, P. I. (1992). Stage-by-stage formation as a method of psychological investigation. *Journal of Russian and East European Psychology*, 30, 60-80.

- Gánem-Gutiérrez, G. A., & Harun, H. (2011). Verbalisation as a mediational tool for understanding tense-aspect marking in English: An application of concept-based instruction. *Language Awareness, 20*(2), 99-119.
- Garcia, P. N. (2012). *Verbalizing in the second language classroom: The development of the grammatical concept of aspect* (Unpublished doctoral dissertation). University of Massachusetts Amherst, Amherst, MA.
- Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Lawrence Erlbaum.
- Grabe, W. (2001). Notes toward a theory in second language writing. In T. Silva & P. K. Matsuda (Eds.), *On second language writing* (pp. 39-57). Mahwah, NJ: Lawrence Erlbaum.
- Grabe, W. (2003). Reading-writing relations: L2 perspectives on research and practice. In B. Kroll (Ed.), *Research on second language writing* (pp. 242-262). New York: Cambridge University Press.
- Graves, D. H. (1994). *A fresh look at writing*. Portsmouth, NH: Heinemann.
- Guinda, C. S. (2012). Proximal positioning in students' graph commentaries. In K. Hyland & C. S. Guinda (Eds.), *Stance and voice in written academic genres* (pp. 166-183). New York: Palgrave Macmillan.
- Hall, G. (2007). Stylistics in second language contexts: A critical perspective. In G. Watson & S. Zyngier (Eds.), *Literature and stylistics for language learners* (pp. 3-14). Hampshire, UK: Palgrave Macmillan.
- Hall, G. (2014). Pedagogical stylistics. In M. Burke (Ed.), *The Routledge handbook of stylistics*, (pp.239-252). New York: Routledge.

- Halliday, M. A. K. (1985). *Introduction to functional grammar* (4th ed.). London: Arnold.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (1999). *Construing experience through meaning*. London: Cassell.
- Hanauer, D. I. (1997). Poetry reading in the second language classroom. *Language Awareness, 6*(1), 2-16.
- Hanauer, D. I. (2001a). The task of poetry reading and second language learning. *Applied Linguistics, 22*(3), 295-323.
- Hanauer, D. I. (2001b). Focus-on-cultural understanding: Literary reading in the second language classroom. *CAUSE: Revista de Filología y su Didáctica, 24*, 389-404.
- Hanauer, D. I. (2010). *Poetry as research: Exploring second language poetry writing*. Amsterdam: John Benjamins.
- Hanauer, D. I. (2011). The scientific study of poetic writing. *Scientific Study of Literature, 1*(1), 79-87.
- Harding, J. R. (2014). Reader response criticism and stylistics. In M. Burke (Ed.), *The Routledge handbook of stylistics* (pp.68-84). New York: Routledge.
- Hegelheimer, V. (2006). Helping ESL writers through a multimodal, corpus-based, online grammar resource. *CALICO Journal, 24*(1), 5-31.
- Hegelheimer, V., & Fisher, D. (2006). Grammar, writing, and technology: A sample technology-supported approach to teaching grammar and improving writing for ESL learners. *CALICO journal, 23*(2), 257-279.
- Helms-Park, R., & Stapleton, P. (2003). Questioning the importance of individualized voice in undergraduate L2 argumentative writing: An empirical study with pedagogical implications. *Journal of Second Language Writing, 12*(3), 245-265.

- Herriman, J. (2013). Systemic functional linguistics. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Oxford, UK: Wiley-Blackwell. doi: 10.1002/9781405198431.wbeal1137
- Hewings, A. (2012). Stance and voice in academic discourse across channels. In K. Hyland & C. S. Guinda (Eds.), *Stance and voice in written academic genres* (pp. 187-201). New York: Palgrave Macmillan.
- Hirvela, A. (2001). Connecting reading and writing through literature. In D. Belcher & A. Hirvela (Eds.), *Linking literacies: Perspectives on L2 reading-writing connections* (pp. 109-134). Ann Arbor, MI: University of Michigan Press.
- Hirvela, A. (2004). *Connecting reading & writing in second language writing instruction*. Ann Arbor, MI: University of Michigan Press.
- Hirvela, A. (2005). ESL students and the use of literature in composition courses. *Teaching English in the Two-Year College*, 33, 70-77.
- Hirvela, A. & Belcher, D. (2001). Coming back to voice: The multiple voices and identities of mature multilingual writers. *Journal of Second Language Writing*, 10, 83-106.
- Hood, S. (2012). Voice and stance as APPRAISAL: Persuading and positioning in research writing across intellectual fields. In K. Hyland & C. S. Guinda (Eds.), *Stance and voice in written academic genres* (pp. 51-68). New York: Palgrave Macmillan.
- Hyland K. (1994). Hedging in academic writing and EAP textbooks. *English for Specific Purposes*, 13, 239-256.
- Hyland, K. (1996a). Writing without conviction? Hedges in science research articles. *Applied Linguistics*, 17, 433-454.

- Hyland, K. (1996b). Hedges, boosters and lexical invisibility: Noticing modifiers in academic texts. *Language Awareness*, 9, 179-197.
- Hyland, K. (1999). Talking to students: Metadiscourse in introductory coursebooks. *English for Specific Purposes*, 18, 3-26.
- Hyland, K. (2004a). *Disciplinary discourses: Social interactions in academic writing*. Ann Arbor: The University of Michigan Press.
- Hyland, K. (2004b). Patterns of engagement: Dialogic features and L2 undergraduate writing. In L. J. Ravelli & R. A. Ellis (Eds.), *Analysing academic writing: contextualized frameworks* (pp. 5-23). London: Continuum.
- Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse studies*, 7(2), 173-192.
- Hyland, K. (2008). Disciplinary voices: Interactions in research writing. *English Text Construction*, 1(1), 5-22.
- Hyland, K. (2012). Undergraduate understandings: Stance and voice in final year reports. In K. Hyland & C. S. Guinda (Eds.), *Stance and voice in written academic genres* (pp. 134-150). New York: Palgrave Macmillan.
- Hyland, K., & Hyland, F. (Eds.). (2006). *Feedback in second language writing: Contexts and issues*. Cambridge, UK: Cambridge University Press.
- Hyland, K., & Guinda, C. S. (Eds.). (2012). *Stance and voice in written academic genres*. New York: Palgrave Macmillan.
- Iida, A. (2010). Developing voice by composing Haiku: A social-expressivist approach for teaching Haiku writing in EFL contexts. *English Teaching Forum*, 48(1), 28-34.



- Iida, A. (2012). The value of poetry writing: Cross-genre literacy development in a second language. *Scientific Study of Literature*, 2(1), 60-82.
- Iida, A. (2013). Critical review of literary reading and writing in a second language. *The Journal of Literature in Language Teaching*, 2, 5-11.
- Ioannidis, J. P. A. (2008). Why most discovered true associations are inflated. *Epidemiology*, 19, 640-648.
- Ivanič, R. & Camps, D. (2001). I am how I sound: Voice as self-representation in L2 writing. *Journal of Second Language Writing*, 10(1-2), 3-33.
- Janssen, van Waes & van den Bergh. (1996). Effects of thinking aloud on writing processes. In L. C. Michael & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 233-250). Mahwah, NJ: Lawrence Erlbaum.
- Jeffery, J. (2009). Constructs of writing proficiency in US state and national writing assessments: Exploring variability. *Assessing Writing*, 14(1), 3-24.
- Jeffries, L. (2000). Don't throw out the baby with the bathwater: in defence of theoretical eclecticism in stylistics. Poetics and Linguistics Association, Occasional Papers, No. 12. Retrieved from: <http://www.pala.ac.uk/uploads/2/5/1/0/25105678/paper12.pdf>
- Johnson, B., & Christensen, L. (2008). *Educational research: Quantitative, qualitative, and mixed approaches*. Thousand Oaks, CA: Sage.
- Johnson, N. H. (2008). *Genre as concept in second language academic writing pedagogy* (Unpublished doctoral dissertation). University of Arizona, Tucson, AZ.
- Jörg, T. 2009. Thinking in complexity about learning and education: A programmatic view. *Complicity: An International Journal of Complexity and Education*, 6, 1-22.

- Jwa, S. (2012). Modeling L2 writer voice: Discoursal positioning in fanfiction writing. *Computers and Composition*, 29(4), 323-340.
- Karpov, Y. V. (2003). Vygotsky's theory of scientific concepts: Its role for contemporary education. In A. Kozulin, B. Gindis, V. S. Ageyev & S. M. Miller (Eds.), *Vygotsky's educational theory in cultural context* (pp. 81-101). Cambridge: Cambridge University Press.
- Kaplan, R. B. (1966). Cultural thought patterns in intercultural education. *Language Learning*, 16, 1-20.
- Kellerman, E. (1985). If at first you do succeed. . . . In S. M. Gass & C. G. Maddens (Eds.), *Input in second language acquisition* (pp. 345–353). Rowley, MA: Newbury House.
- Khatib, M., Rezaei, S., & Derakhshan, A. (2011). Literature in EFL/ESL classroom. *English Language Teaching*, 4(1), 201-208.
- Kim, J. (2013). *Developing conceptual understanding of sarcasm in a second language through concept-based instruction* (Unpublished doctoral dissertation). The Pennsylvania State University, University Park, PA.
- Kim, M. (2004). Literature discussions in adult L2 learning. *Language and Education*, 18(2), 145-166.
- Kim, T-Y., Baba, K. & Cumming, A. (2006). Goals, motivations, and identities of three students writing in English. In A. Cumming (Ed.), *Goals for academic writing: ESL students and their instructors* (pp.125-141). Amsterdam: John Benjamins.
- Kirshner, D. & Kellogg, D. 2009. Lev Vygotsky as muse to complex learning/teaching: A response to Ton Jörg's programmatic view. *Complicity: An International Journal of Complexity and Education*, 6, 45-55.

- Kline, R. B. (2004). *Beyond significance testing: Reforming data analysis methods in behavioral research*. Washington, DC: American Psychological Association.
- Retrieved from:  
<http://edhar.cas.msu.edu/tc985/readings/Kline%20Significance%20-%20ch2%20-%20refresher%20-%20hires.pdf>
- Knouzi, I., Swain, M., Lapkin, S., & Brooks, L. (2010). Self-scaffolding mediated by languaging: Microgenetic analysis of high and low performers. *International Journal of Applied Linguistics*, 20(1), 23-49.
- Kobayashi, H. (1984). *Rhetorical patterns in English and Japanese* (Unpublished doctoral dissertation). Columbia University Teachers College, New York, NY.
- Kramsch, C. (2000). Social discursive constructions of self in L2 learning. In J. P. Lantolf (Ed.), *Sociocultural Theory and Second Language Learning* (pp. 133-153). Oxford: Oxford University Press.
- Kubota, R. (1992). *Contrastive rhetoric of Japanese and English: A critical approach* (Unpublished doctoral dissertation). University of Toronto, Toronto, Canada.
- Kubota, R. (1997). Reevaluation of the uniqueness of Japanese written discourse: Implications for contrastive rhetoric. *Written Communication*, 14(4), 460-480.
- Kubota, R. (1998). An investigation of L1-L2 transfer in writing among Japanese university students: Implications for contrastive rhetoric. *Journal of Second Language Writing*, 7(1), 69-100.
- Kuo, C. H. (1999). The use of personal pronouns: Role relationships in scientific journal articles. *English for Specific Purposes*, 18(2), 121-138.

- Lai, W. (2012). *Concept-based foreign language pedagogy: Teaching the Chinese temporal system* (Unpublished doctoral dissertation). The Pennsylvania State University, University Park, PA.
- Lam, W.S.E. (2000). L2 Literacy and the design of the self: A case study of a teenager writing on the Internet. *TESOL Quarterly*, 34(3), 457-482.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174.
- Lantolf, J. P., & Poehner, M. E. (2007). Dynamic assessment of L2 development: Bringing the past into the future. *Journal of Applied Linguistics and Professional Practice*, 1(1), 49-72.
- Lantolf, J. P., & Poehner, M. E. (2008). Introduction to sociocultural theory and the teach of second languages. In J.P. Lantolf & M. E. Poehner (Eds.), *Sociocultural theory and the teaching of second languages* (pp. 1-32). London: Equinox.
- Lantolf, J. P., & Poehner, M. E. (2014). *Sociocultural theory and the pedagogical imperative in L2 education: Vygotskian praxis and the research/practice divide*. New York: Routledge.
- Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford: Oxford University Press.
- Lapkin, S., Swain, M., & Knouzi, I. (2008). Postsecondary French as a second language students learn the grammatical concept of voice: Study design, materials development, and pilot data. In J. P. Lantolf & M. E. Poehner (Eds.), *Sociocultural theory and the teaching of second languages* (pp. 228-255). London: Equinox.

- Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, 27(4), 590-619.
- Larsen-Freeman, D. (2011). A complexity theory approach to second language development/acquisition. In D. Atkinson (Ed.), *Alternative approaches to second language acquisition* (pp. 48-72). New York: Routledge.
- Larsen-Freeman, D. (2012). On the roles of repetition in language teaching and learning. *Applied Linguistics Review*, 3(2), 195-210.
- Larsen-Freeman, D. (2013). Chaos/complexity theory for second language acquisition. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Oxford, UK: Wiley- Blackwell. doi: 10.1002/9781405198431.wbeal0125
- Larsen-Freeman, D. & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford: Oxford University Press.
- Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. New York: Routledge.
- Lavelli, M., Pantoja, A. P., Hsu, H., Messinger, D., & Fogel, A. (2005). Using microgenetic designs to study change processes. In D. M. Teti (Ed.), *Handbook of research methods in developmental science* (pp. 40-65). Malden, NJ: Wiley- Blackwell. doi:10.1002/9780470756676.ch3
- Lee, H. (2012). *Concept-based approach to second language teaching and learning: Cognitive linguistics-inspired instruction of English phrasal verbs* (Unpublished doctoral dissertation). The Pennsylvania State University, University Park, PA.

- Leech, G. & Short, M. (2007). *Style in fiction: A linguistic introduction to English fictional prose* (2nd ed.). Harlow, UK: Longman.
- Leki, I. (2007). *Undergraduates in a second language. Challenges and complexities of academic literacy development*. New York: Lawrence Erlbaum.
- Leki, I., Cumming, A., & Silva, T. (2008). *A synthesis of research on second language writing in English*. London: Routledge.
- Leech, G. & Short, M. (2007). *Style in fiction: A linguistic introduction to English fictional prose* (2nd ed.). Harlow, UK: Longman.
- Lightbown, P. M. (1983). Exploring relationships between developmental and instructional sequences in L2 acquisition. In H. Seliger & M. H. Long (Eds.), *Classroom oriented research in second language acquisition* (pp. 217-243). Rowley, MA: Newbury House.
- Lightbown, P. M. (2000). Anniversary article. Classroom SLA research and second language teaching. *Applied Linguistics*, 21(4), 431-4462.
- Lin, H. (2010). The taming of the immeasurable: An empirical assessment of language awareness. In A. Paran & L. Sercu (Eds.), *Testing the untestable in language education* (pp. 191-216). Toronto: Multilingual Matters
- Liu, P. L. (2011). A study on the use of computerized concept mapping to assist ESL learners' writing. *Computers & Education*, 57(4), 2548-2558.
- Lobman, C. (2010). Creating developmental moments: Teaching and learning as creative activities. In V.P. John-Steiner, M. C. Connery & A. Marjanovic-Shane (Eds.), *Vygotsky and creativity: A cultural-historical approach to play, meaning making, and the arts* (pp. 199-214). New York: Peter Lang.

- Manchón, R. M. (2009). *Writing in foreign language contexts: Learning, teaching, and research*. Bristol, UK: Multilingual Matters.
- Manchón, R. M., Roca de Larios, J., & Murphy, L. (2009). The temporal dimension and problem-solving nature of foreign language composing processes: Implications for theory. In R. M. Manchón (Ed.), *Writing in foreign language contexts: Learning, teaching, and research* (pp. 102-129). Bristol, UK: Multilingual Matters.
- Marsen, S. (2012). Detecting the creative in written discourse. *Writing and Pedagogy*, 4(2), 209-231.
- Matsuda, P. K. (2001). Voice in Japanese written discourse: Implications for second language writing. *Journal of Second Language Writing*, 10(1-2), 35-53.
- Matsuda, P. K. & Jeffery, J. V. (2012). Voice in student essays. In K. Hyland & C. S. Guinda (Eds.), *Stance and voice in written academic genres*. (pp. 151-165). New York: Palgrave Macmillan.
- Matsuda, P. K., & Tardy, C. M. (2007). Voice in academic writing: The rhetorical construction of author identity in blind manuscript review. *English for Specific Purposes*, 26(2), 235-249.
- McKay, S. (2001). Literature as content for ESL/EFL. In M. Celce-Murcia (Ed.), *Teaching English as a Second or Foreign Language* (3rd ed., pp. 319-332). Boston: Heinle & Heinle.
- McNamara, D.S., Crossley, S.A. & McCarthy, P.M. (2010). The linguistic features of quality writing. *Written Communication*, 27(1), 57-86.
- McRae, J. (1997). *Literature with a small "l"* (2nd ed.). London: Macmillan/Prentice-Hall.

- Miles, M.B., & Huberman, A.M. (1994). *Qualitative data analysis* (2nd ed.). Thousand Oaks, CA: Sage.
- Myskow, G., & Gordon, K. (2012). Getting interpersonal on a university entrance exam impromptu writing task. *Journal of English for Academic Purposes*, 11(2), 90-98.
- Negueruela, E. (2003). *A sociocultural approach to the teaching and learning of second languages: Systemic-theoretical instruction and L2 development* (Unpublished doctoral dissertation). The Pennsylvania State University, University Park, PA.
- Negueruela, E. & J.P. Lantolf. (2005). Concept-based instruction: Teaching grammar in an intermediate-advanced Spanish L2 university classroom. *CALPER Working Papers Series*, No. 3. The Pennsylvania State University, Center for Advanced Language Proficiency Education and Research, University Park, PA. Retrieved from: <http://calper.la.psu.edu/publication.php?page=wps3>
- Ojima, M. (2006). Concept mapping as pre-task planning: A case study of 3 Japanese ESL writers. *System*, 34, 566-585.
- Ong, J. (2014). How do planning time and task conditions affect metacognitive processes of L2 writers? *Journal of Second Language Writing*, 23, 17-30.
- O'Sullivan, I. (2007). Enhancing a process-oriented approach to literacy and language learning: The role of corpus consultation literacy. *ReCALL*, 19(3), 269-286.
- Paesani, K. (2006). 'Exercices de style': Developing multiple competencies through a writing portfolio. *Foreign Language Annals*, 39(4), 618-639.
- Papousek, H. & Papousek, M. (1984). Qualitative transitions in integrative processes during the first trimester of human postpartum life. In H. F. R. Prechtel (Ed.), *Continuity of*



- neural functions from prenatal to postnatal life* (pp. 220–41). Oxford: Blackwell Scientific.
- Paran, A. (2008). The role of literature in instructed foreign language learning and teaching: An evidence-based survey. *Language Teaching*, 41(4), 465-496.
- Park, K. (2012). Learner–corpus interaction: A locus of microgenesis in corpus-assisted L2 writing. *Applied Linguistics*, 33(4), 361-385.
- Pavlenko, A. (2007). Autobiographic narratives as data in applied linguistics. *Applied Linguistics*, 28, 163-188.
- Perneger, T. V. (1998). What’s wrong with Bonferroni adjustments. *British Medical Journal*, 316, 1236-1238.
- Plummer, P. & Busse, B. (2006). E-learning and *Language and Style* in Mainz and Münster. *Language and Literature*, 15(3), 257-276.
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878-912.
- Prior, P. (2001). Voice in text, mind, and society: Sociohistoric accounts of discourse acquisition and use. *Journal of Second Language Writing*, 10(1–2), 55-81.
- Ramanathan, V., & Kaplan, R. B. (1996). Audience and voice in current L1 composition texts: Some implications for ESL student writers. *Journal of Second Language Writing*, 5(1), 21-34.
- Ramanathan, V. & Atkinson, D. (1999). Individualism, academic writing, and ESL writers. *Journal of Second Language Writing*, 8(1), 45-75.

- Reid, J. (2001). Advanced EAP writing and curriculum design: What do we need to know?  
In T. Silva & P. K. Matsuda (Eds.), *On second language writing* (pp. 143-160).  
Mahwah, NJ: Lawrence Erlbaum.
- Roca de Larios, J., Manchón, R. M, Murphy, L. & Marín, J. (2008). The foreign language  
writer's strategic behaviour in the allocation of time to writing processes. *Journal of  
Second Language Writing, 17*, 30-47.
- Rosenblatt, L. M. (1978). *The reader, the text, the poem*. Carbondale, IL: Southern Illinois  
University Press.
- Rosenthal, J. A. (1996). Qualitative descriptors of strength of association and effect size.  
*Journal of Social Service Research, 21*(4), 37-59.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research* (2nd ed.). Newbury Park,  
CA: Sage.
- Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges  
(Eds.), *The handbook of research synthesis* (pp. 231-244). New York: Sage.
- Salekin, R. T., Rogers, R., & Sewell, K. W. (1996). A review and meta-analysis of the  
psychopathy checklist and psychopathy checklist-revised: Predictive validity of  
dangerousness. *Clinical Psychology: Science and Practice, 3*(3), 203-215.
- Sasaki, M. (2004). A multiple-data analysis of the 3.5-year development of EFL student  
writers. *Language Learning, 54*, 525-582.
- Saugera, V. (2011). Scriptwriting as a tool for learning stylistic variation. *Foreign Language  
Annals, 44*(1), 137-152.
- Shen, F. (1989). The classroom and the wider culture: Identity as a key to learning English  
composition. *College Composition and Communication, 40*(4), 459-466.

- Sheskin, D. J. (2003). *Handbook of parametric and nonparametric statistical procedures*. Boca Raton, FL: Chemical Rubber Company Press.
- Short, M. (1996). *Exploring the language of poems, plays and prose*. London: Pearson Education.
- Siegler, R.S. (1995). How does change occur: A microgenetic study of number conservation. *Cognitive Psychology*, 25, 225-273.
- Siegler, R.S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.
- Siegler, R. S. (2006). Microgenetic analyses of learning. In D. Kuhn & R. S. Siegler (Eds.), *Handbook of child psychology, volume 2: Cognition, perception, and language* (6th ed., pp. 464–510). Hoboken, NJ: Wiley.
- Siegler, R. S., & Svetina, M. (2002). A microgenetic/cross-sectional study of matrix completion: Comparing short-term and long-term change. *Child Development*, 73, 793–809.
- Silva, T., & Matsuda, P. K. (Eds.). (2010). *Practicing theory in second language writing*. West Lafayette, IN: Parlor Press.
- Simpson, P. (1997). *Language through literature: An introduction*. London: Routledge.
- Slate, J., & Rojas-LeBouef, A. (2011). *Calculating basic statistical procedures in SPSS: A self-help and practical guide to preparing theses, dissertations, and manuscripts*. Retrieved from <http://cnx.org/content/col11292/1.6/>
- Smagorinsky, P. (1989). The reliability and validity of protocol analysis. *Written Communication*, 6(4), 463-479.

- Smagorinsky, P. (1992). How reading model essays affects writing. In J. Irwin & M. A. Doyle (Eds.), *Reading/writing connections: Learning from research* (pp. 160-176). Newark, DE: International Reading Association.
- Sperling, M. & Appleman, D. (2011). Voice in the context of literacy studies. *Reading Research Quarterly*, 46, 70-84.
- Spiro, J. (2010). Crossing the bridge from appreciative reader to reflective writer: The assessment of creative process. In A. Paran & L. Sercu (Eds.), *Testing the untestable in language education* (pp. 165-190). Bristol: Multilingual Matters.
- Spoelman, M. & Verspoor, M. (2010). Dynamic patterns in development of accuracy and complexity: A longitudinal case study in the acquisition of Finnish. *Applied Linguistics*, 31, 532-553.
- Storch, N., & Wigglesworth, G. (2010). Learners' processing, uptake, and retention of corrective feedback on writing. *Studies in Second Language Acquisition*, 32(2), 303-334.
- Strauss, A. L. & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage.
- Swain, M. (2006). Languaging, agency and collaboration in advanced second language learning. In H. Byrnes (Ed.), *Advanced language learning: The contributions of Halliday and Vygotsky* (pp. 95-108). London: Continuum.
- Swain, M. (2010). "Talking-it-through": Languaging as a source of learning. In R. Batstone (Ed.), *Sociocognitive perspectives on second language learning and use* (pp. 112-129). Oxford: Oxford University Press.

- Swain, M., Kinnear, P. & Steinman, L. (2010). *Sociocultural theory in second language education*. Bristol, UK: Multilingual matters.
- Swain, M., & Lapkin, S. (2002). Talking it through: Two French immersion learners' response to reformulation. *International Journal of Educational Research*, 37, 285-304.
- Swain, M., Lapkin, S., Knouzi, I., Suzuki, W. and Brooks, L. (2009). Languaging: University students learn the grammatical concept of voice in French. *Modern Language Journal*, 93, 5-29.
- Swain, M. & Watanabe, Y. (2013). Languaging: Collaborative dialogue as a source of second language learning. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Oxford: Wiley-Blackwell. doi: 10.1002/9781405198431.wbeal0664
- Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Tardy, C. M. (2009). *Building genre knowledge*. West Lafayette, IN: Parlor Press.
- Thompson, P. (2012). Achieving a voice of authority in PhD theses. In K. Hyland & C. S. Guinda (Eds.), *Stance and voice in written academic genres*. (pp. 119-133). New York: Palgrave Macmillan.
- Thompson, G. & Ye, Y. (1991). Evaluation of the reporting verbs used in academic papers. *Applied Linguistics* 12, 365-382.
- Timucin, M. (2001). Gaining insights into alternative teaching approaches employed in an EFL literature class. *CAUCE, Revista de Filología y su Didáctica*, 24, 269-293.
- Todd, R. W., Khongputb, S., & Darasawanga, P. (2007). Coherence, cohesion and comments on students' academic essays. *Assessing Writing*, 12, 10-25.

- Trevarthen, C. (1977). Descriptive analysis of infant communicative behavior. In H. R. Schaffer (Ed.), *Studies in mother–infant interaction* (pp. 227–70). New York: Academic Press.
- Tse, P. (2012). Stance in academic bios. In K. Hyland & C. S. Guinda (Eds.), *Stance and voice in written academic genres*. (pp. 69-84). New York: Palgrave Macmillan.
- Twyman, T., Ketterlin-Geller, L. R., McCoy, J. D., & Tindal, G. (2003). Effects of concept-based instruction on an English language learner in a rural school: A descriptive case study. *Bilingual Research Journal*, 27(2), 259-274.
- U2 (1983). Sunday bloody Sunday. On *War*. London: Island Records.
- van Compernelle, R. A. (2011). Developing second language sociopragmatic knowledge through concept-based instruction: A microgenetic case study. *Journal of Pragmatics*, 43(13), 3267-3283.
- van Compernelle, R. A. (2012). *Developing sociopragmatic capacity in a second language through concept-based instruction* (Unpublished doctoral dissertation). The Pennsylvania State University, University Park, PA.
- van den Bergh, H. & Rijlaarsdam, G. C. (2001). Changes in cognitive activities during the writing process and relations with text quality. *Educational Psychology*, 21, 373-385.
- Vandrick, S. (2003). Literature in the teaching of second language composition. In B. Kroll (ed.), *Exploring the dynamics of second language writing* (pp.263-283). Cambridge, UK: Cambridge University Press.
- van Geert, P. 1998. A dynamic systems model of basic developmental mechanisms: Piaget, Vygotsky and beyond. *Psychological Review*, 105, 634-677.

- van Horn, R., Jr. (2000). Banished to my room. Retrieved from  
<http://www.storybytes.com/view-stories/2000/banished-to-my-room.html>
- van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. (1994). *The think aloud method: A practical guide to modelling cognitive processes* (Vol. 2). London: Academic Press.
- Verdonk, P. (2002). *Stylistics*. Oxford: Oxford University Press.
- Verdonk, P. (2013). *The stylistics of poetry: Context, cognition, discourse, history*. London: Bloomsbury.
- Verspoor, M., Lowie, W. & van Dijk, M. (2008). Variability in second language development from a dynamic systems perspective. *The Modern Language Journal* 92, 214-231.
- Verspoor, M., Schmid, M., & Xu, X. (2012). A dynamic usage based perspective on L2 writing. *Journal of Second Language Writing*, 21, 3, 239-263.
- Vygotsky, L. S. (1978) *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Vygotsky, L. S. (1986). *Thought and language* (Revised ed.). Cambridge, MA: MIT Press.
- Vygotsky, L. S. (1987). *The collected works of L. S. Vygotsky. Volume 1: Problems of general psychology, including the volume Thinking and Speech*. (N. Minick, Trans.). R. W. Rieber & A. S. Carton (Eds.). New York: Plenum.
- Wales, K. (2006). Stylistics. In *Encyclopedia of language and linguistics*. (2nd ed., pp. 213-217). Amsterdam: Elsevier.
- Wang, P. L. (2009). The application of integrated literature instruction in freshman English teaching. *Sino-US English Teaching*, 6(9), 1-11.

- Warner, C. (2012). Literary pragmatics in the advanced foreign language literature classroom: The case of *Young Werther*. In M. Burke, S. Csábi, L. Week & J. Zerkowitz (Eds.), *Pedagogical stylistics: Current trends in language, literature and ELT* (pp.142-157). London: Continuum.
- Warschauer, M. (2002). Networking into academic discourse. *Journal of English for Academic Purposes*, 1(1), 45-58.
- Watanabe, Y., & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners. *Language Teaching Research*, 11, 121-142.
- Watson, P. (2013). *Rules of thumb on magnitudes of effect sizes*. Retrieved from <http://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/effectSize>
- Weir, I. (n. d.). *Spearman's rank correlation—introduction*. Retrieved from <http://www.statstutor.ac.uk/resources/uploaded/spearmans.pdf>
- Weissberg, R. (2006). *Connecting speaking and writing in second language writing instruction*. Ann Arbor, MI: University of Michigan.
- Wertsch, J. (1991). *Voices of the mind: A socio-cultural approach to mediated action*. Cambridge, MA: Harvard University Press.
- Wertsch, J. V. & Hickmann, M. (1987). Problem solving in social interaction: A microgenetic analysis. In M. Hickmann (Ed.), *Social and functional approaches to language and thought* (pp. 251-266). Orlando, FL: Academic Press.
- Wertsch, J. V. & Stone, C. A. (1978). Microgenesis as a tool for developmental analysis. *Laboratory of Comparative Human Cognition*, 1, 8-10.
- Widdowson, H.G. (1975). *Stylistics and the teaching of literature*. London: Longman.



- Wu, S. & Rubin, D. (2000). Evaluating the impact of collectivism and individualism on argumentative writing by Chinese and North American college students. *Research in the Teaching of English*, 35, 148-178.
- Yañez-Prieto, M.C. (2008). *On literature and the secret art of (im)possible worlds: Teaching literature through language* (Unpublished doctoral dissertation). The Pennsylvania State University, University Park, PA.
- Yañez-Prieto, M.C. (2010). Authentic instruction in literary worlds: Learning the stylistics of concept-based grammar. *Language and Literature*, 19(1), 59-75.
- Yetkin Özdemir, E. I. (2011). Self-regulated learning from a sociocultural perspective. *Egitim Ve Bilim*, 36(160), 298. Retrieved from <http://search.proquest.com/docview/1009841835?accountid=14771>
- Zerkowitz, J. (2012). Stylistics for language teachers. In M. Burke, S. Csábi, L. Week & J. Zerkowitz (Eds.), *Pedagogical stylistics: Current trends in language, literature and ELT* (pp.193-209). London: Continuum.
- Zhao, C. G. (2010). *The role of voice in high-stakes second language writing assessment: Teaching literature through language* (Unpublished doctoral dissertation). New York University, New York, NY.
- Zhao, C. G. (2013). Measuring authorial voice strength in L2 argumentative writing: The development and validation of an analytic rubric. *Language Testing*, 30(2), 201-230.
- Zyngier, S., & Fialho, O. (2010). Pedagogical stylistics, literary awareness and empowerment: a critical perspective. *Language and Literature*, 19(1), 13-33.

Zyngier, S., Fialho, O., & Rios, P. A. P. (2007). Revisiting literary awareness: In G. Watson, & S. Zyngier (Eds.), *Literature and stylistics for language learners: Theory and practice* (pp. 194–209). New York: Palgrave Macmillan.

## Appendices

### Appendix A Recruitment Poster

# FREE TOEFL iBT Essay-Writing Course!!

Small class size - 8 students only!

Starting in September, 2013

Would you like to improve your TOEFL iBT® writing skills?

Are you a **Japanese** student studying English?

Are you 18-26 years old?

Have you studied at a university in Japan? **Or** are you planning to study at a university in Canada?

If you answered YES to ALL of these questions you may qualify to participate in a research study that offers a FREE 3-week TOEFL iBT essay-writing course.

If you are interested, please contact the email address below for more information.

-----@gmail.com

**Appendix B**  
Information Letter and Consent Form for Outside Raters

Dear potential research participant,

My name is Gary Fogal. I am a Ph.D. student at the Ontario Institute for Studies in Education (OISE), at the University of Toronto. I am writing to ask for your participation in research I am doing for my thesis research. After reading the detailed information below, if you wish to participate in this study please complete and return to me via email the consent form attached to the bottom of this letter. Thank you.

Title of Research Project: *Pedagogical stylistics and concept-based instruction: An investigation into the development of voice in the academic writing of Japanese university students of English*

Principal Investigator: Gary G. Fogal, Ph.D. Candidate, Department of Curriculum, Teaching & Learning, OISE/University of Toronto

Purpose of the Study: This study aims to examine how a teaching aid called pedagogical stylistics can help Japanese university students improve the quality of voice in their academic writing, specifically with respect to the TOEFL iBT independent writing task.

Participants: I am looking to recruit three English language instructors for this study. In order to participate in this study you must:

- 1) be trained professionally to be an English language instructor (for example, you should possess a TESOL Certificate or a CELTA Certificate, or the like), or you should have an advanced degree in the field of second language acquisition (for example, an MA in Applied Linguistics or TESOL, or the like); and
- 2) have at least 3 years of experience scoring essays associated with standardized proficiency tests, such as the TOEFL or the IELTS.

Benefits: Your participation in this study will enhance your understanding of the role of voice in academic writing, specifically for the TOEFL iBT independent writing task. You will also be exposed to a variety of writing styles that may serve your own professional needs. Your participation in a rater training session will also expose you to a variety of scoring schemes and/or scoring techniques that may be new to you. This may also serve your professional interests.

What Participants will do: During this study you will be asked to score anonymous essays based on writing prompts typically found on the TOEFL iBT independent writing task. Each

essay will be approximately 300 words long. Two separate scoring rubrics will be provided for you. You will be asked to score each essay using the TOEFL iBT independent writing task scoring rubric and a voice scoring rubric. In order to train you with using these rubrics and to ensure rater agreement with other raters, you will be asked to participate in a rater training session that will take approximately 2 hours. This session is tentatively scheduled for the afternoon of September 3, 2013. In total, you will be asked to score 78 essays over 3 months (approximately 6.5 essays per week). You will also receive a \$150 honorarium for your time as a participant in this study.

### Participants' Rights

- **To Confidentiality:** I will not collect any data from you other than the scores you assign to each essay. I will reference the essays you score using a pseudonym. Your name will not be used in the reporting of any data in scholarly publications or conferences.
- **To Ask Questions about the Research:** If you wish to ask questions about this research project, you may do so at any point. Please contact me (Gary Fogal) at xxx-xxx-xxxx or gg.fogal@xxxx.ca, or you may speak to me directly. You may also contact my supervisor, Alister Cumming, regarding questions at xxx-xxx-xxxx or alister.cumming@xxxx.ca. The University of Toronto also has an office regarding ethics if you want more information about your rights as a research participant, or to verify the authenticity of this research. You may contact the Office of Research Ethics at xxx-xxx-xxxx or ethics.review@xxxx.ca.
- **To Withdraw at Any Time:** You may withdraw from this study at any time. If you do withdraw before your role in the study is completed, you can indicate to me in writing whether you wish me to use or immediately to destroy the scores you provided. Your \$150 honorarium will also be pro-rated based on the number of essays you have scored.

**Risks:** There are no potential risks to you should you decide to participate in this study.

Please read and complete the attached consent form to indicate your willingness, or not, to participate in this study.

Kindly,

Gary G. Fogal  
OISE / University of Toronto

PLEASE KEEP A COPY OF THIS LETTER FOR YOUR RECORDS

\*\*\*\*\*

### Consent Form for Participants

I have read the letter from Gary Fogal that describes my role in the project, *Pedagogical stylistics and concept-based instruction: An investigation into the development of voice in the academic writing of Japanese university students of English*. I understand that all of the data I provide will be kept confidential and secure, and the data will be destroyed after a period of four years. I also understand that I may withdraw from the study at any time, with the only consequence that my \$150 honorarium will be pro-rated. I also understand that the data collected from me may be used in academic journals or academic presentations. In such cases, I understand that no identifying information will be presented about me. I also acknowledge that my questions regarding this study have been answered satisfactorily. As such,

- I agree
- I do not agree

to participate in this study.

Name (print): \_\_\_\_\_

Email or Phone Number: \_\_\_\_\_

Signature: \_\_\_\_\_

Date: \_\_\_\_\_

**Appendix C**  
Biodata Questionnaire

(administered online)

1. Pseudonym (Fake Name): \_\_\_\_\_
2. Gender: \_\_\_\_\_ Male      \_\_\_\_\_ Female
3. Age: \_\_\_\_\_
4. What is/are your first language(s): \_\_\_\_\_
5. Beside your first language(s) and English, do you know any other languages? Yes\_ No \_\_  
If yes, which languages?

\_\_\_\_\_

How well do you know them (beginner, intermediate, advanced)?

\_\_\_\_\_

6. What was your major area of study (your major) at university?

\_\_\_\_\_

7. In Japan, have you ever studied argumentative or persuasive essay writing in Japanese?  
Yes\_\_ No \_\_

If you remember, can you provide more details about this (for example, at what age, the teaching style, the specific topics of study, etc):

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

8. In Japan, have ever studied Japanese literature in school, in Japanese? Yes\_\_ No \_\_

If yes, can you briefly describe how studied this topic (for example, what kind of exercises did your teachers ask you to do?):

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

9. How old were you when you started to study English? \_\_\_\_\_

10. How many years have you studied English? \_\_\_\_\_

11. Please complete the following table:

<i>Where have you studied English? (Tick as many as needed.)</i>	<i>In what country did you study at these places? (You may include more than one country if necessary.)</i>	<i>For how long did you study at each place? (weeks, months or years)</i>
<input type="checkbox"/> High School		
<input type="checkbox"/> University		
<input type="checkbox"/> Language School		
<input type="checkbox"/> Private Lessons		

12. Did you speak English at home or outside of school when you were growing up?

Yes\_\_ No \_\_

If yes, can you provide more details about this (for example, with whom, how often, the age you started, etc):

---



---



---

13. Regarding your life these days, how many days every week do you do the following:

	<i>For example: 2 days, 5 days, 0 days</i>
Study English formally in a school or with a tutor.	
Study English formally outside of school (for example, doing homework or studying for a test).	
Read in English for entertainment (Facebook, magazines, etc).	
Write in English for entertainment (Facebook, diary entries, etc).	

14. Are you taking any English classes at present, or will you be taking any English classes at the same time as the course you are taking with me? If so, please list these classes below.

<i>School Name</i>	<i>Frequency (Daily, Once a week, etc.)</i>	<i>Class Name or Focus (Grammar, Tourism English, Academic Writing, Conversation, etc.)</i>
_____	_____	_____
_____	_____	_____
_____	_____	_____



15. Why are you participating in this study?

---



---



---

16. What standardized English proficiency tests have you taken (TOEFL, TOEIC, IELTS, etc.), when did you take the test(s), and if you recall, what was your score(s)? If you have taken any of these tests more than once, please report each time you took the test separately.

<i>Test Name</i>	<i>Date</i>	<i>Score</i>	<i>Score on the Writing Section (if you remember)</i>
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____

17. Why are you studying for the TOEFL test? (What are your future education or employment plans?)

---



---



---

18. If you are planning to take the TOEFL test in the future, when do you plan to take it?

October – December 2013	_____
January – March 2014	_____
April – June 2014	_____
July - September 2014	_____
October December 2014	_____
Other?	_____

**Appendix D**  
TOEFL iBT Independent Writing Task Rubric (Education Testing Service, 2008)

 **TOEFL® iBT Test**  
**Independent Writing Rubrics (Scoring Standards)**

Score	Task Description
5	<p><b>An essay at this level largely accomplishes all of the following:</b></p> <ul style="list-style-type: none"> <li>• effectively addresses the topic and task</li> <li>• is well organized and well developed, using clearly appropriate explanations, exemplifications, and/or details</li> <li>• displays unity, progression, and coherence</li> <li>• displays consistent facility in the use of language, demonstrating syntactic variety, appropriate word choice, and idiomaticity, though it may have minor lexical or grammatical errors</li> </ul>
4	<p><b>An essay at this level largely accomplishes all of the following:</b></p> <ul style="list-style-type: none"> <li>• addresses the topic and task well, though some points may not be fully elaborated</li> <li>• is generally well organized and well developed, using appropriate and sufficient explanations, exemplifications, and/or details</li> <li>• displays unity, progression, and coherence, though it may contain occasional redundancy, digression, or unclear connections</li> <li>• displays facility in the use of language, demonstrating syntactic variety and range of vocabulary, though it will probably have occasional noticeable minor errors in structure, word form, or use of idiomatic language that do not interfere with meaning</li> </ul>
3	<p><b>An essay at this level is marked by one or more of the following:</b></p> <ul style="list-style-type: none"> <li>• addresses the topic and task using somewhat developed explanations, exemplifications, and/or details</li> <li>• displays unity, progression, and coherence, though connection of ideas may be occasionally obscured</li> <li>• may demonstrate inconsistent facility in sentence formation and word choice that may result in lack of clarity and occasionally obscure meaning</li> <li>• may display accurate but limited range of syntactic structures and vocabulary</li> </ul>
2	<p><b>An essay at this level may reveal one or more of the following weaknesses:</b></p> <ul style="list-style-type: none"> <li>• limited development in response to the topic and task</li> <li>• inadequate organization or connection of ideas</li> <li>• inappropriate or insufficient exemplifications, explanations, or details to support or illustrate generalizations in response to the task</li> <li>• a noticeably inappropriate choice of words or word forms</li> <li>• an accumulation of errors in sentence structure and/or usage</li> </ul>
1	<p><b>An essay at this level is seriously flawed by one or more of the following weaknesses:</b></p> <ul style="list-style-type: none"> <li>• serious disorganization or underdevelopment</li> <li>• little or no detail, or irrelevant specifics, or questionable responsiveness to the task</li> <li>• serious and frequent errors in sentence structure or usage</li> </ul>
0	<p><b>An essay at this level merely copies words from the topic, rejects the topic, or is otherwise not connected to the topic, is written in a foreign language, consists of keystroke characters, or is blank.</b></p>

Copyright © 2008 by Educational Testing Service. All rights reserved.

TOEFL iBT Independent Writing Task Rubric. From “TOEFL iBT Test: Independent Writing Rubrics (Scoring Standards),” by Educational Testing Service, 2008. Reprinted by permission of Educational Testing Service, the copyright owner. No endorsement of any kind by Educational Testing Service should be inferred.

**Appendix E**  
Voice Rubric

Rhetorical feature of voice	Score	Your Score	Score Descriptors
Hedges	5		<ul style="list-style-type: none"> <li>-The writer presents ideas and claims with definitive language that shows authority and confidence.</li> <li>-The reader feels that the writer has a clear stance on and a strong attitude toward the topic under discussion.</li> <li>-The tone of the writing shows personality, adds life to the writing, and is engaging and appropriate for the intended reader.</li> <li>-Word choice, and language use by extension, is varied, often interesting, sophisticated, and eye-catching to the reader.</li> </ul>
	4		
	3		<ul style="list-style-type: none"> <li>-The writer presents ideas and claims somewhat mildly with frequent use of unnecessary hedges; only occasionally does the writing show some degree of authority and confidence.</li> <li>-The writer seems to have a stance on the topic under discussion, but no strong attitude is revealed in the writing.</li> <li>-The tone of the writing is appropriate for the intended reader and the purpose of the writing, but lacks personality and liveliness.</li> <li>-Occasional interesting word choice and language use may catch the reader's attention, but the effect is inconsistent.</li> </ul>
	2		
	1		<ul style="list-style-type: none"> <li>-The writer presents ideas and claims very mildly, showing a lack of authority and confidence in what he/she is writing.</li> <li>-The writer seems indifferent and does not have a clear stance on or attitude toward the topic under discussion.</li> <li>-The writer writes in a monotone that does not engage the reader at all; oftentimes the reader find him- or herself drifting off while reading the text.</li> <li>-Word choice or language use is flat, general, and dull, and thus unable to catch the reader's attention.</li> </ul>

Boosters	5	<ul style="list-style-type: none"> <li>-The writer presents ideas and claims with definitive language that shows authority and confidence.</li> <li>-The reader feels that the writer has a clear stance on and a strong attitude toward the topic under discussion.</li> <li>-The tone of the writing shows personality, adds life to the writing, and is engaging and appropriate for the intended reader.</li> <li>-Word choice, and language use by extension, is varied, often interesting, sophisticated, and eye-catching to the reader.</li> </ul>
	4	
	3	<ul style="list-style-type: none"> <li>-The writer presents ideas and claims somewhat mildly with frequent use of unnecessary hedges; only occasionally does the writing show some degree of authority and confidence.</li> <li>-The writer seems to have a stance on the topic under discussion, but no strong attitude is revealed in the writing.</li> <li>-The tone of the writing is appropriate for the intended reader and the purpose of the writing, but lacks personality and liveliness.</li> <li>-Occasional interesting word choice and language use may catch the reader's attention, but the effect is inconsistent.</li> </ul>
	2	
	1	<ul style="list-style-type: none"> <li>-The writer presents ideas and claims very mildly, showing a lack of authority and confidence in what he/she is writing.</li> <li>-The writer seems indifferent and does not have a clear stance on or attitude toward the topic under discussion.</li> <li>-The writer writes in a monotone that does not engage the reader at all; oftentimes the reader find him- or herself drifting off while reading the text.</li> <li>-Word choice or language use is flat, general, and dull, and thus unable to catch the reader's attention.</li> </ul>
attitude markers	5	<ul style="list-style-type: none"> <li>-The writer presents ideas and claims with definitive language that shows authority and confidence.</li> <li>-The reader feels that the writer has a clear stance on and a strong attitude toward the topic under discussion.</li> <li>-The tone of the writing shows personality, adds life to the writing, and is engaging and appropriate for the intended reader.</li> </ul>

		-Word choice, and language use by extension, is varied, often interesting, sophisticated, and eye-catching to the reader.
	4	
	3	-The writer presents ideas and claims somewhat mildly with frequent use of unnecessary hedges; only occasionally does the writing show some degree of authority and confidence. -The writer seems to have a stance on the topic under discussion, but no strong attitude is revealed in the writing. -The tone of the writing is appropriate for the intended reader and the purpose of the writing, but lacks personality and liveliness. -Occasional interesting word choice and language use may catch the reader's attention, but the effect is inconsistent.
	2	
	1	-The writer presents ideas and claims very mildly, showing a lack of authority and confidence in what he/she is writing. -The writer seems indifferent and does not have a clear stance on or attitude toward the topic under discussion. -The writer writes in a monotone that does not engage the reader at all; oftentimes the reader find him- or herself drifting off while reading the text. -Word choice or language use is flat, general, and dull, and thus unable to catch the reader's attention.
authorial self-mention	5	-The writer reveals him- or herself in the writing either directly or indirectly, giving the reader a clear sense of who the writer is as a unique individual. -The reader feels that the writer is aware of and able to engage the reader effectively in either a direct or a subtle way. -The sharing of personal backgrounds and experiences, if any, is effective, genuine, and engaging to the reader.
	4	
	3	-The writer reveals him- or herself in the writing to some extent, leaving the reader with some sense of who he/she is. -The reader feels that the writer is aware of and trying to engage the reader in a way, but with limited success. -The sharing of personal backgrounds and experiences, if any, is genuine but not so engaging or

		effective to the reader.
	2	
	1	<ul style="list-style-type: none"> <li>-The reader has little or no sense of who the writer is as a unique individual instead of a generic, faceless person.</li> <li>-The reader feels that the writer is not concerned with the reader or completely fails to engage the reader in any way.</li> <li>-The sharing of personal backgrounds and experiences, if any, is generic, ineffective, and even inappropriate, making the reader feel annoyed.</li> </ul>
direct reader reference	5	<ul style="list-style-type: none"> <li>-The writer reveals him- or herself in the writing either directly or indirectly, giving the reader a clear sense of who the writer is as a unique individual.</li> <li>-The reader feels that the writer is aware of and able to engage the reader effectively in either a direct or a subtle way.</li> <li>-The sharing of personal backgrounds and experiences, if any, is effective, genuine, and engaging to the reader.</li> </ul>
	4	
	3	<ul style="list-style-type: none"> <li>-The writer reveals him- or herself in the writing to some extent, leaving the reader with some sense of who he/she is.</li> <li>-The reader feels that the writer is aware of and trying to engage the reader in a way, but with limited success.</li> <li>-The sharing of personal backgrounds and experiences, if any, is genuine but not so engaging or effective to the reader.</li> </ul>
	2	
	1	<ul style="list-style-type: none"> <li>-The reader has little or no sense of who the writer is as a unique individual instead of a generic, faceless person.</li> <li>-The reader feels that the writer is not concerned with the reader or completely fails to engage the reader in any way.</li> <li>-The sharing of personal backgrounds and experiences, if any, is generic, ineffective, and even inappropriate, making the reader feel annoyed.</li> </ul>

Based on the above, please select one from below. Indicate your selection here: \_\_\_\_\_

Overall Voice Strength	
5	The reader feels a very strong authorial voice in the writing that is compelling and engaging.
4	The reader feels a fairly strong authorial voice in the writing.
3	The reader feels a somewhat weak authorial voice in the writing.
2	The reader feels a very weak authorial voice in the writing.
1	The reader cannot really feel the presence of an authorial voice in the writing.

Voice rubric. Adapted from “Measuring authorial voice strength in L2 argumentative writing: The development and validation of an analytic rubric,” by C. G. Zhao, 2013, *Language Testing, Volume 30*, pp. 226-229. Copyright 2013 by Sage. Adapted with permission.

### Appendix F

#### Tally Framework for Measuring Frequency of Use of Rhetorical Features

Participant ID: \_\_\_\_\_

TOEFL writing task ID: \_\_\_\_\_

<i>Rhetorical feature of voice</i>	<i>Frequency</i>
hedges	
boosters	
attitude markers	
authorial self-mention	
direct reader reference	



**Appendix G**  
Stylistic Analysis Worksheet

Identification

Code: \_\_\_\_\_

Name: \_\_\_\_\_

Date: \_\_\_\_\_

I. Detect and Identify

1. Which of the following stylistic/rhetorical features did you find in today's text?

- Tick (✓) the box if you think the feature is *present* in the text.
- Leave the box empty if you think the feature is *absent* from the text.

- A. hedge
- B. booster
- C. attitude marker
- D. direct reader reference
- E. authorial/narrator self-mention

2. Based on what you have just ticked (✓), indicate where in the text the feature is by circling the feature and indicating next to the circle the corresponding letter. For example:

“They might<sup>A</sup> be coming down for tea,” Frances says.

## II. Describe

In your opinion, what might be the function(s) of the rhetorical feature(s) you found in the text? Try to comment on all of the rhetorical features you ticked (✓) above.

Stylistic/rhetorical features	How does the feature contribute meaning or significance to the text?

## III. Discuss

In a group, discuss some of the features you identified in part II., and how the feature(s) contribute meaning or significance to the text. If possible, offer your group members your own example of this feature.

## Appendix H

### Script for Combined Stimulated Recall & Semi-structured Interview Protocol

Hello, and thank you for agreeing to sit-down with me to discuss some of your writing and the course you took with me. This interview is going to be divided into two parts. I'll talk about Part One now, and then I'll talk about Part Two later, after we finish Part One. I expect this interview to take about 45 minutes, so thank you in advance for your time.

#### I. PART ONE

For part one, we're going to be talking about the essay you wrote on DATE. I have a copy of it here. Please take a few minutes now to reread it to refresh your memory.

A. I'm now going to ask you a few questions about your writing:

1. In paragraph X, sentence Y, you did Z. Can you explain why you made this choice and what effect you hoped it would have on the reader?
2. Is there anything that we studied in class or discussed in class that may have influenced this choice? Please be as specific as possible.

*(Question 1 and 2 will be repeated until I have exhausted the items I wish to enquire about. I suspect there will be about 5-7 items I will enquire about.)*

B. I'm now going to ask you general questions about this piece of writing:

1. What were your goals for completing this writing task?
2. Do you feel like you have accomplished these goals?
3. Overall, but with respect to voice in particular, how would you describe the way you presented yourself in your writing? If possible, please provide an example from your writing to illustrate your comments.

C. I'm now going to give you the opportunity to answer two less constrictive questions, meaning that you should feel free to comment about whatever you wish or however you wish about these two questions.

1. First, would you like to add anything about this piece of writing that I have not asked about?
2. Second, would you like to add anything about your writing in general that I have not asked about?

Thank you. We are now going to move on to Part Two of this interview.

## II. PART TWO

Part Two will focus on your experiences in the TOEFL writing course you just finished with me, instead of just talking specifically about one writing task. This part of the interview is a semi-structured interview. This means that I do have a few questions prepared for you, but I may ask you follow-up questions based on your responses. This also means that if you want to talk about something I haven't asked you about, please feel free to do so. Do you have any questions about the format of this part of the interview?

With all of the questions I'm about to ask you, and when possible, please try and be as specific as possible with your commentary, or provide examples if you can.

1. Can you talk about any specific changes, positive or negative, to your writing that you think are a result from having taken this course?
2. Did you find anything we studied or discussed in class particularly helpful for your overall English academic writing ability?
3. In general, what do you think about using stylistics to learn about voice in academic writing?
4. Would you recommend studying writing with the help of stylistics to others? Why or why not?
5. In the first few days of this course, I gave you a diagram – the OBA – to work with. Here it is again to refresh your memory. Can you comment on whether or not it was helpful, and if so, how?
6. With the exception of the first few days, you were in daily group discussions regarding the stylistic analysis you did for the day. Can you talk about if these discussions influenced your learning of the rhetorical features of voice?
7. Can you also talk about if these discussions influenced your academic writing?
8. Finally, please feel free to comment on anything you wish with respect to the TOEFL course, or anything we discussed today.

Thank you kindly for your time. I really appreciate you having participated in this study.

## Appendix I

### TOEFL iBT Independent Writing Task Essay Prompts

#### *I. Essay prompts informing the baseline data*

1. Do you agree or disagree with the following statement? Universities should give the same amount of money to their students' sports activities as they give to their university libraries. Use specific reasons and examples to support your opinion.
2. Some people prefer to work for a large company. Others prefer to work for a small company. Which would you prefer? Use specific reasons and details to support your choice.
3. Some people like to travel with a companion. Other people prefer to travel alone. Which do you prefer? Use specific reasons and examples to support your choice.

#### *II. Essay prompts informing intervention-phase data*

1. Which would you choose: a high-paying job with long hours that would give you little time with family and friends or a lower-paying job with shorter hours that would give you more time with family and friends? Explain your choice, using specific reasons and details.
2. Do you agree or disagree with the following statement? Only people who earn a lot of money are successful. Use specific reasons and examples to support your answer.
3. Some people say that computers have made life easier and more convenient. Other people say that computers have made life more complex and stressful. What is your opinion? Use specific reasons and examples to support your answer.
4. Do you agree or disagree with the following statement? Watching television is bad for children. Use specific details and examples to support your answer.

#### *III. Essay prompts informing posttest and delayed posttest data*

1. Some people prefer to live in a small town. Others prefer to live in a big city. Which place would you prefer to live in? Use specific reasons and details to support your answer.
2. Do you agree or disagree with the following statement? Face-to-face communication is better than other types of communication, such as letters, email, or telephone calls. Use specific reasons and details to support your answer.

### Appendix J

#### Coding Framework for Concept Development of Rhetorical Features of Voice in TOEFL iBT Independent Writing Task and Stylistic Worksheets

Data Source	Definition	Code
TOEFL writing task & stylistic worksheet	Participant correctly used the conceptual unit outside of the context in which it was learned. Moreover, he/she correctly identified the conceptual unit on the stylistic worksheet. His/her analysis of the conceptual unit also demonstrated an understanding of its function(s).	5
TOEFL writing task & stylistic worksheet	Participant correctly used the conceptual unit outside of the context in which it was learned. Moreover, he/she correctly identified the conceptual unit on the stylistic worksheet. However, he/she did not provide an accompanying explanation on the stylistic worksheet or the accompanying explanation did not clearly explain the function of the conceptual unit.	4
TOEFL writing task & stylistic worksheet	Participant correctly used the conceptual unit outside of the context in which it was learned. However, he/she was unable to identify the conceptual unit on the stylistic worksheet.	3
Stylistic worksheet	Participant correctly identified the conceptual unit. Moreover, his/her analysis of the conceptual unit demonstrated an understanding of its function(s).	2
Stylistic worksheet	Participant correctly identified the conceptual unit. However, he/she did not provide an accompanying explanation of the conceptual unit or the accompanying explanation did not clearly explain the function of the conceptual unit.	1
Stylistic worksheet	Participant failed to identify the conceptual unit.	0

Coding framework for concept development of rhetorical features of voice in TOEFL iBT independent writing task and stylistic worksheets. Adapted from “Mediating between scientific and spontaneous concepts through languaging,” by L. Brooks, M. Swain, and I. Knouzi, 2010, *Language Awareness, Volume 19*, p. 95. Copyright 2010 by Routledge. Adapted with permission.

**Appendix K**  
Coding Framework for Concept Development of Rhetorical Features of Voice in Audio Recordings of Classroom Discussions

Data Source	Definition	Code
Audio recording of group discussion	Participant correctly used the conceptual unit outside of the context in which it was learned. Moreover, he/she correctly identified the conceptual unit in his/her discussion. His/her analysis of the conceptual unit also demonstrated an understanding of its function(s).	5
Audio recording of group discussion	Participant correctly used the conceptual unit outside of the context in which it was learned. Moreover, he/she correctly identified the conceptual unit in his/her discussion. However, he/she did not provide an accompanying explanation of the conceptual unit or the accompanying explanation did not clearly explain the function of the conceptual unit.	4
Audio recording of group discussion	Participant correctly used the conceptual unit outside of the context in which it was learned. However, he/she did not provide an accompanying explanation of the conceptual unit or the accompanying explanation did not clearly explain the function of the conceptual unit.	3
Audio recording of group discussion	Participant correctly identified the conceptual unit. Moreover, his/her analysis of the conceptual unit demonstrated an understanding of its function(s).	2
Audio recording of group discussion	Participant correctly identified the conceptual unit. However, he/she did not provide an accompanying explanation of the conceptual unit or the accompanying explanation did not clearly explain the function of the conceptual unit.	1
Audio recording of group discussion	Participant failed to identify the conceptual unit.	0

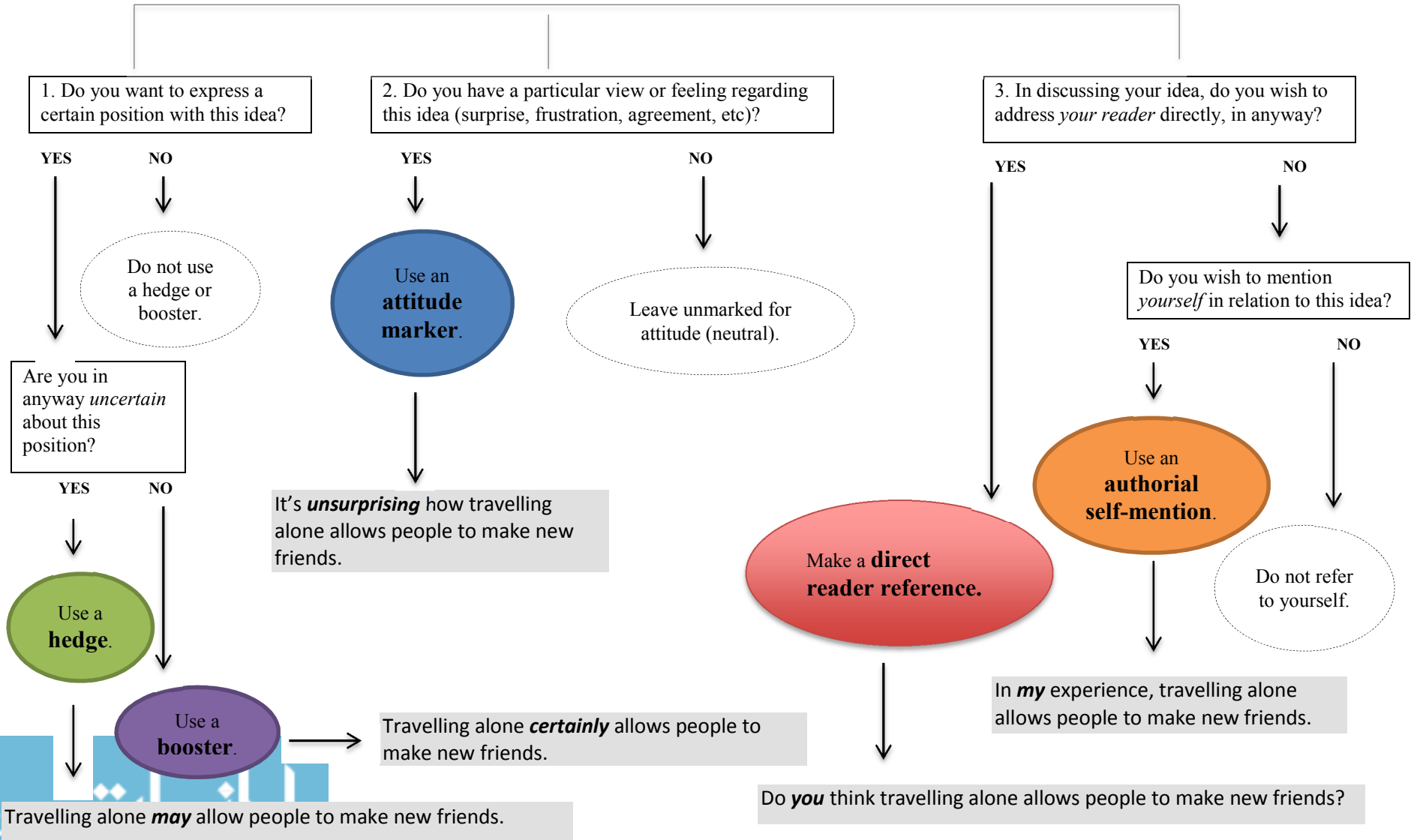
Coding framework for concept development of rhetorical features of voice in audio recordings of classroom discussions. Adapted from “Mediating between scientific and spontaneous concepts through languaging,” by L. Brooks, M. Swain, and I. Knouzi, 2010, *Language Awareness, Volume 19*, p. 95. Copyright 2010 by Routledge. Adapted with permission.

**Appendix L**  
Orienting Basis of Action

Sentence level idea you wish to express.



Eg. Travelling alone allows people to make new friends.





Appendix M  
Visual Aid II

Express yourself with confidence



University students who choose to live alone will **definitely** learn life-long skills.

Uncertainty, or you want to leave room for other interpretations or possibilities



University students choosing to live alone **might** encounter certain advantages.

Express some attitude – share your feeling about a topic



**Frankly**, university students who choose to live alone will learn life-long skills.



Speak directly to your reader

As a university student, how would **you** feel if you had to live with total strangers?

Insert yourself directly into the text



Based on **my own** experiences, **I** think university students would benefit from living alone.

## Appendix N

### Line Graphs for RQ1

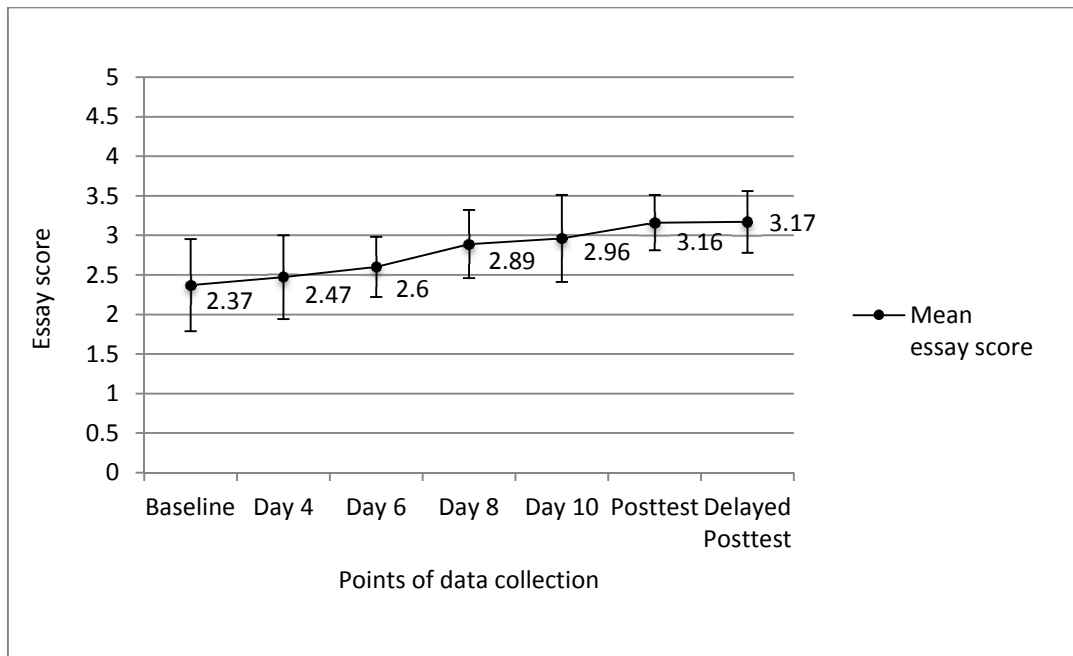


Figure N1. Change over time regarding overall quality of writing ( $n = 7$ ).

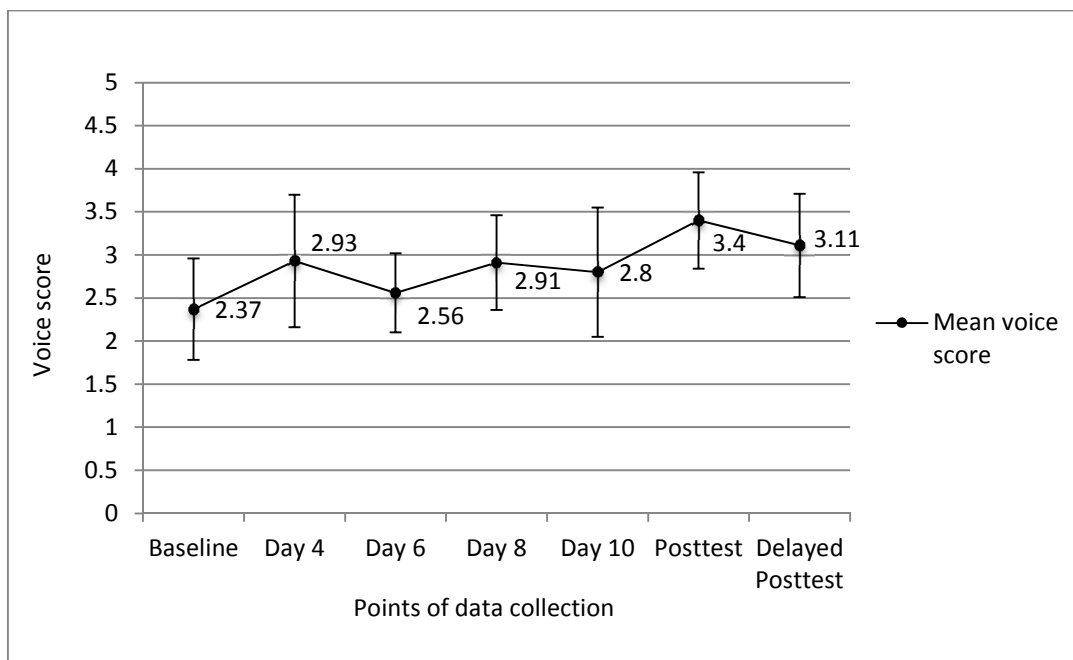
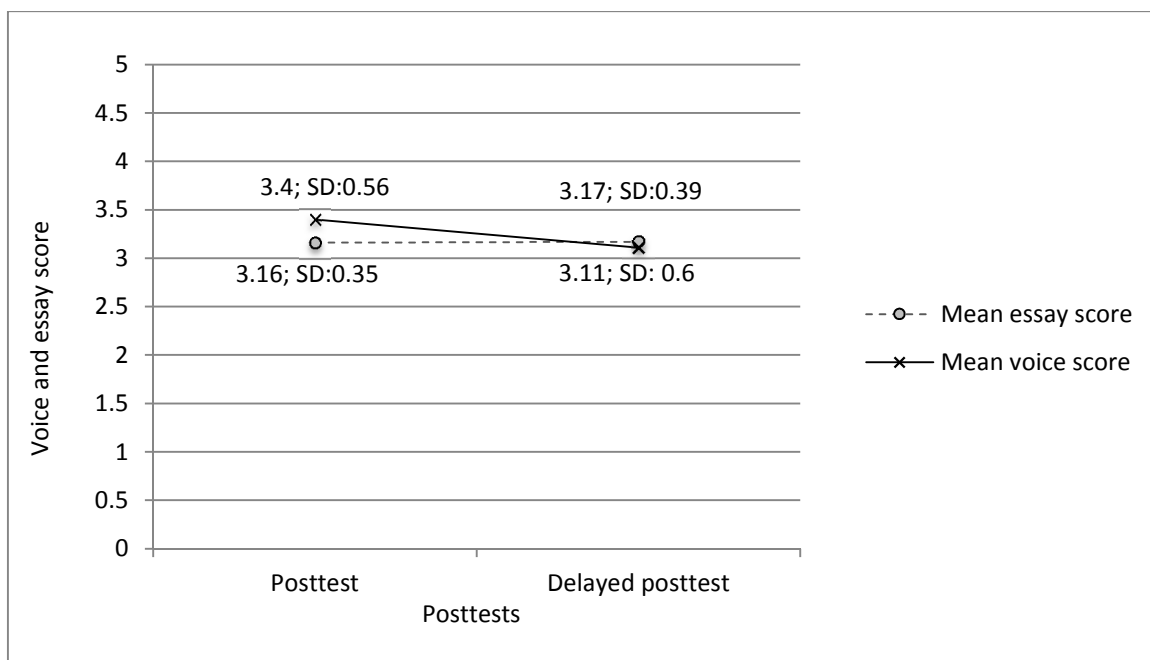


Figure N2. Change over time regarding overall quality of voice ( $n = 7$ ).

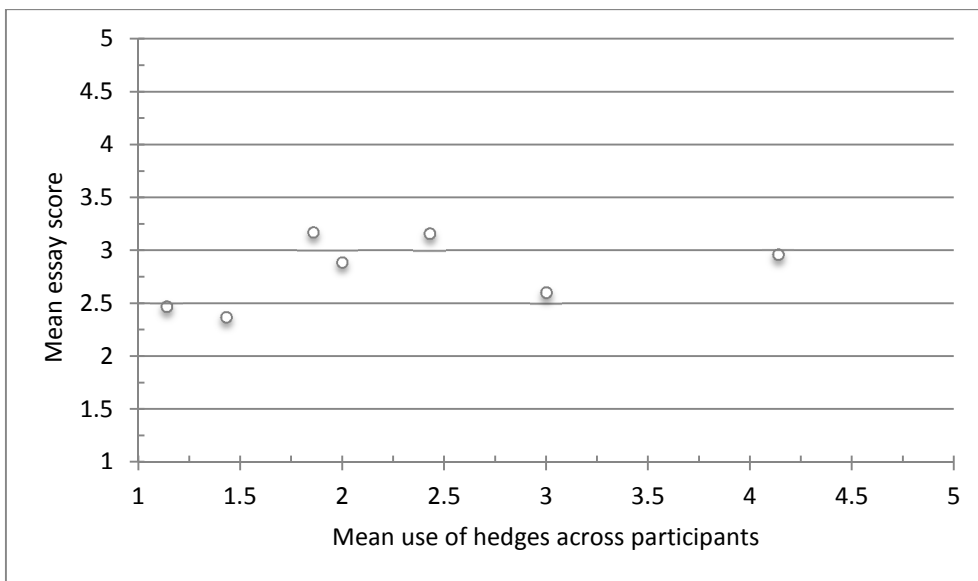
**Appendix O**  
Line Graph for RQ1 (part two)



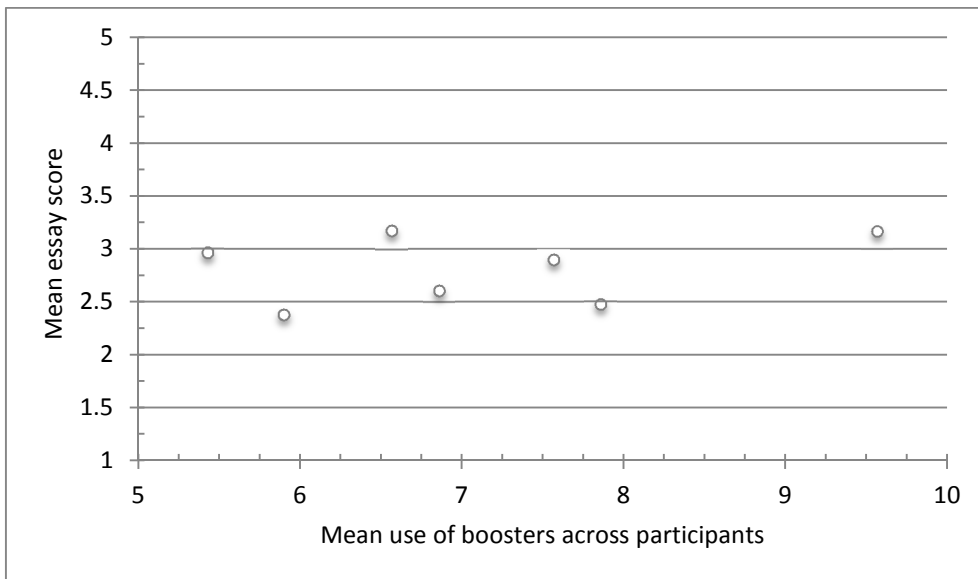
*Figure O1.*

Effects of the intervention sustained over time based on the posttest and the delayed posttest ( $n = 7$ ).

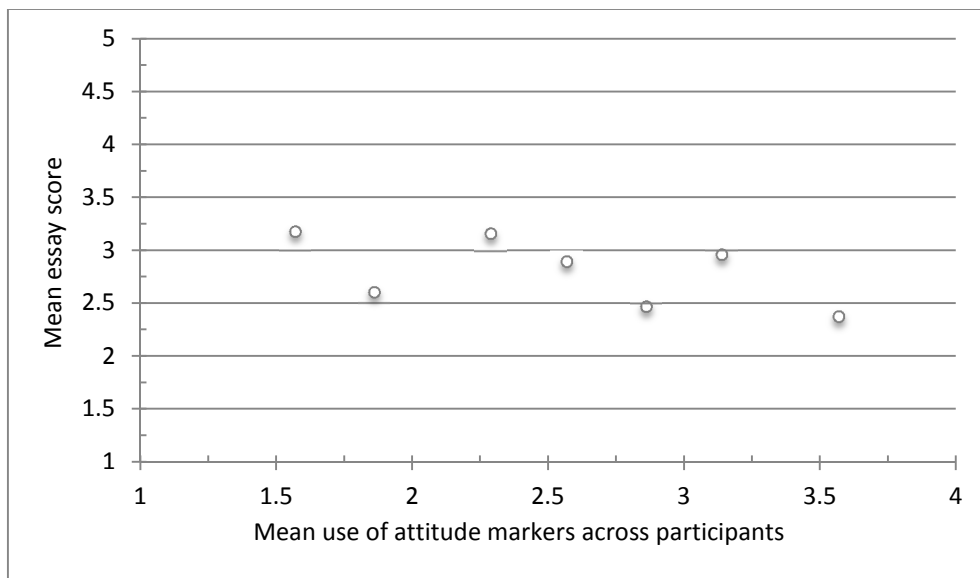
**Appendix P**  
Scatterplots for RQ2: Writing Scores and Rhetorical Feature Use



*Figure P1.*  
Relationship between the frequency of use of hedges and essay scores ( $n = 7$ ).

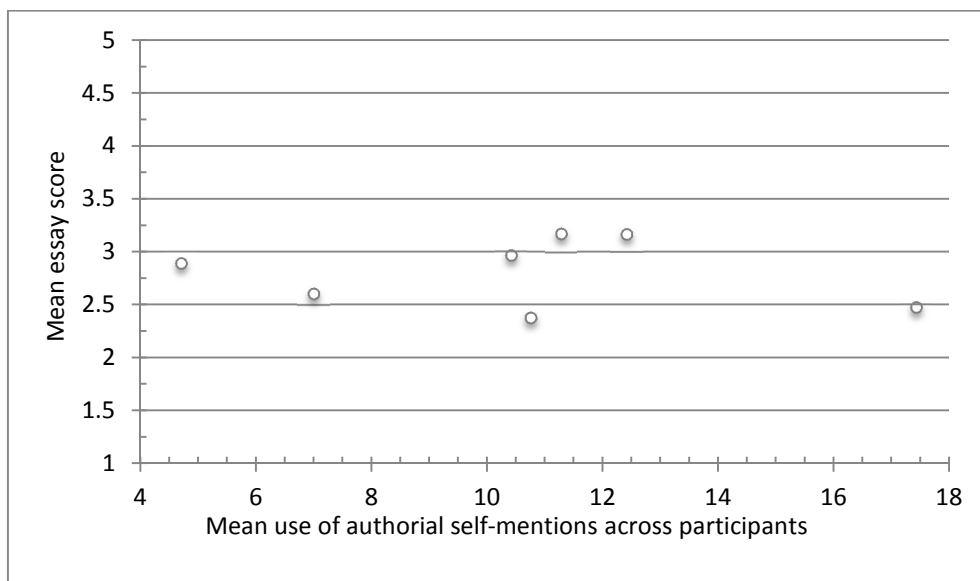


*Figure P2.*  
Relationship between the frequency of use of boosters and essay scores ( $n = 7$ ).



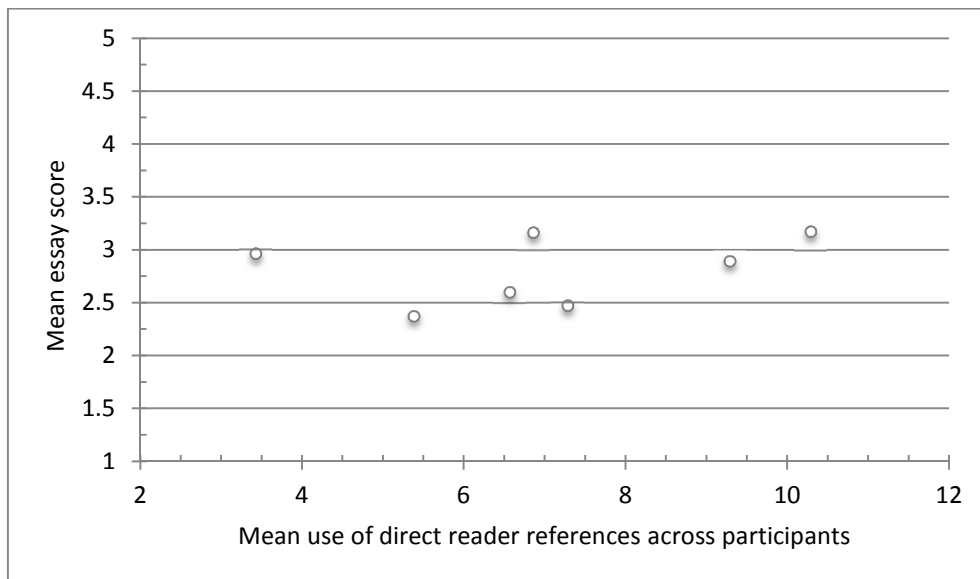
*Figure P3.*

Relationship between the frequency of use of attitude markers and essay scores ( $n = 7$ ).



*Figure P4.*

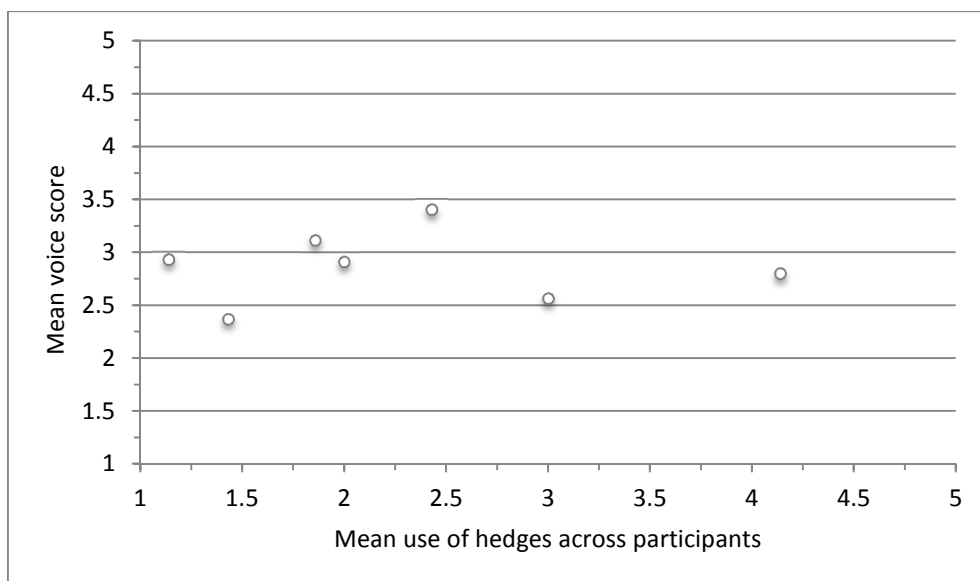
Relationship between the frequency of use of authorial self-mention and essay scores ( $n = 7$ ).



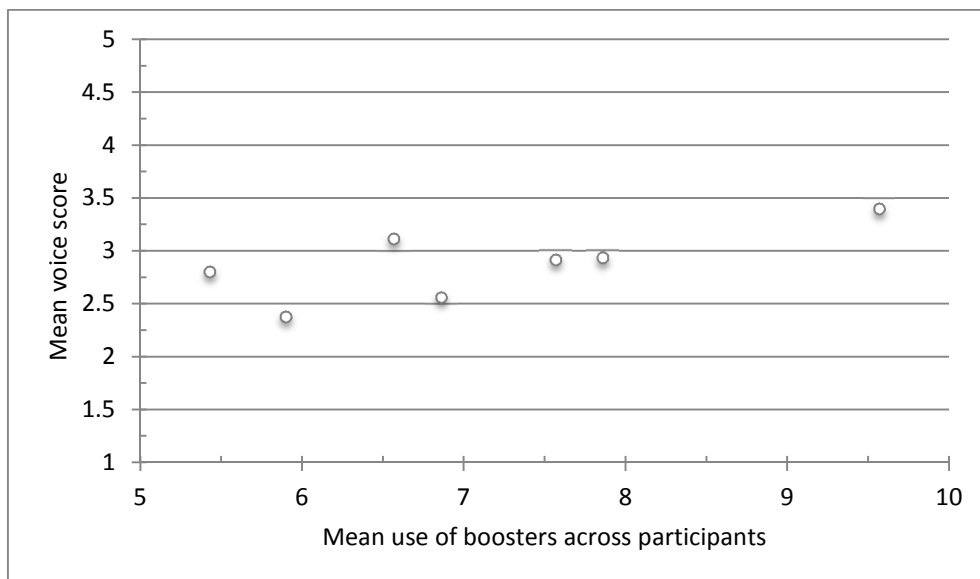
*Figure P5.*

Relationship between the frequency of use of direct reader references and essay scores ( $n = 7$ ).

**Appendix Q**  
Scatterplots for RQ2: Voice Scores and Rhetorical Feature Use



*Figure Q1.*  
Relationship between the frequency of use of hedges and voice scores ( $n = 7$ ).



*Figure Q2.*  
Relationship between the frequency of use of boosters and voice scores ( $n = 7$ ).

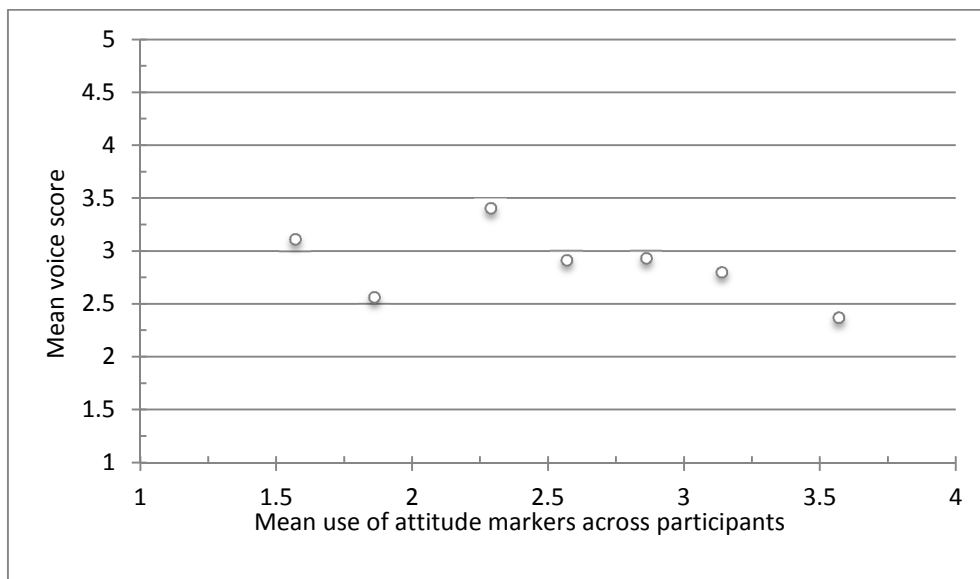


Figure Q3.

Relationship between the frequency of use of attitude markers and voice scores ( $n = 7$ ).

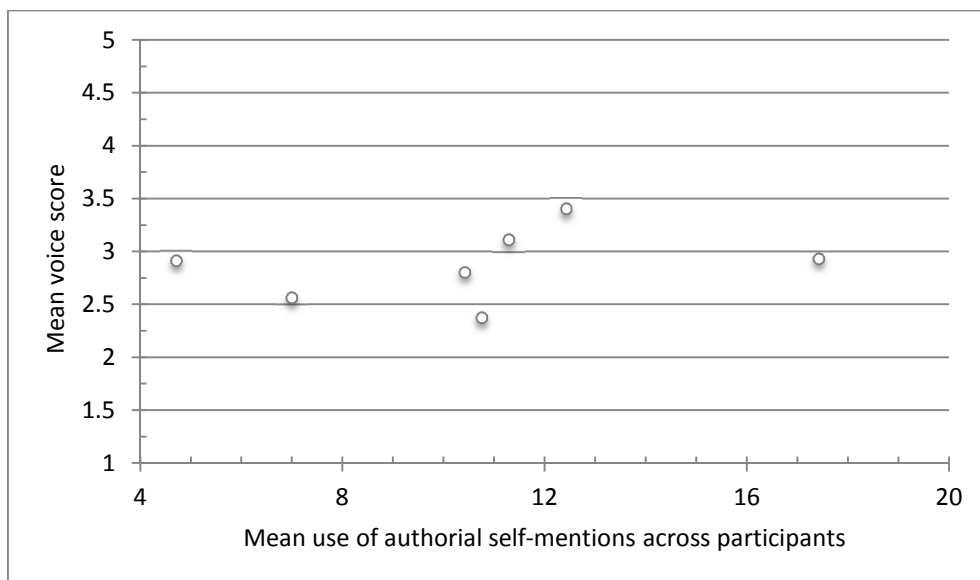


Figure Q4.

Relationship between the frequency of use of authorial self-mention and voice scores ( $n = 7$ ).



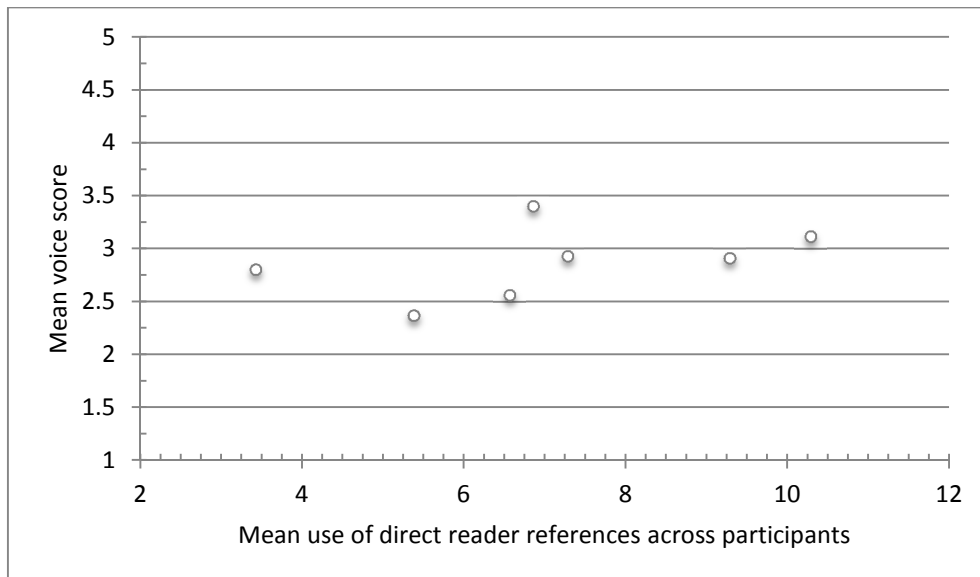
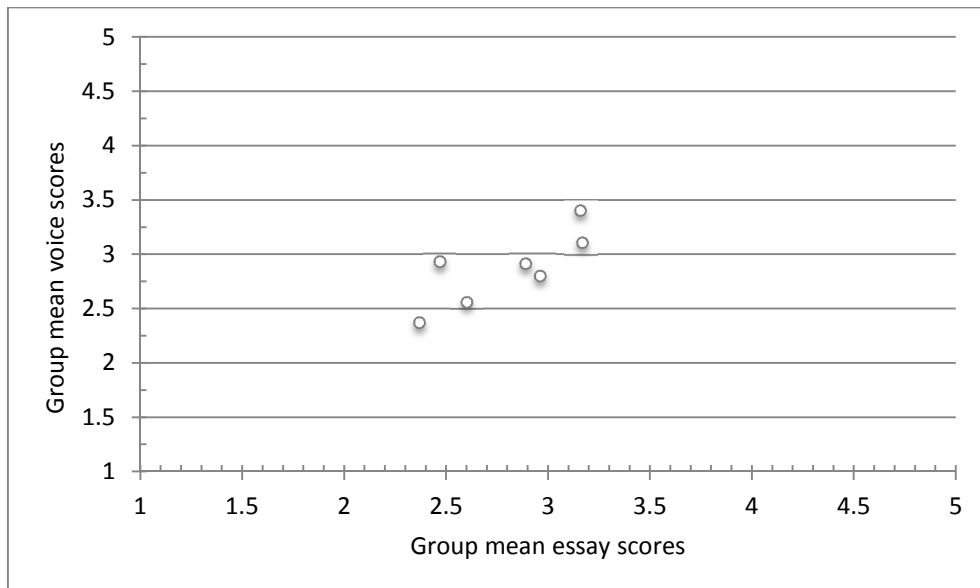


Figure Q5.

Relationship between the frequency of use of direct reader references and voice scores ( $n = 7$ ).

**Appendix R**  
Scatterplot for RQ3: Writing Scores and Voice Scores

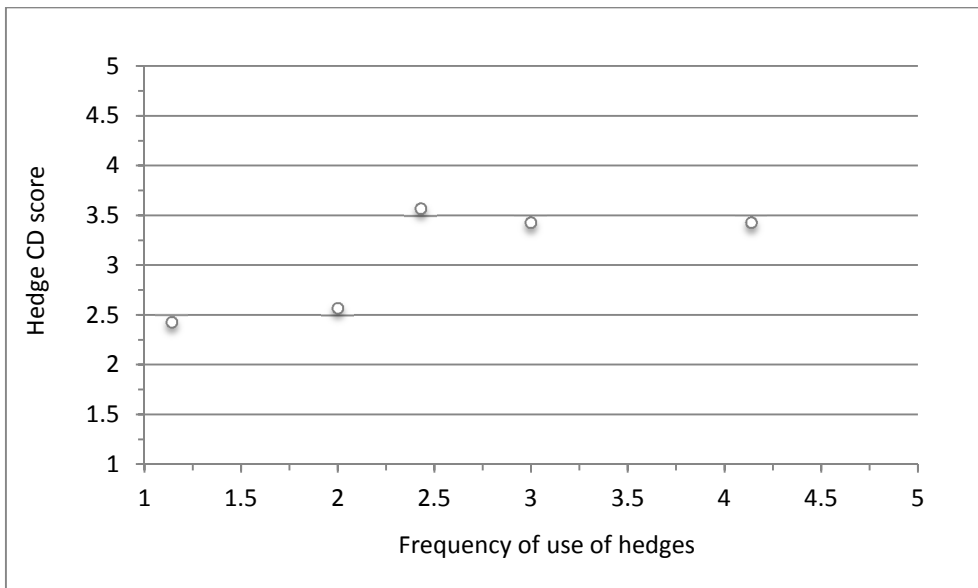


*Figure R1.*

Relationship between writing scores and voice scores across writing tasks and across participants ( $n = 7$ ).

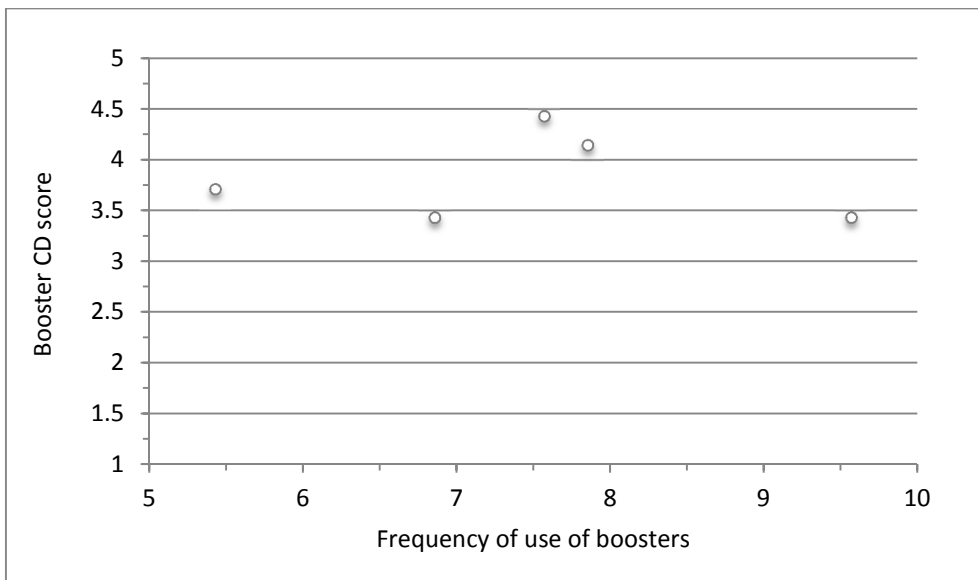
### Appendix S

#### Scatterplots for RQ4: Rhetorical Feature Concept Development Scores and Writing Tasks



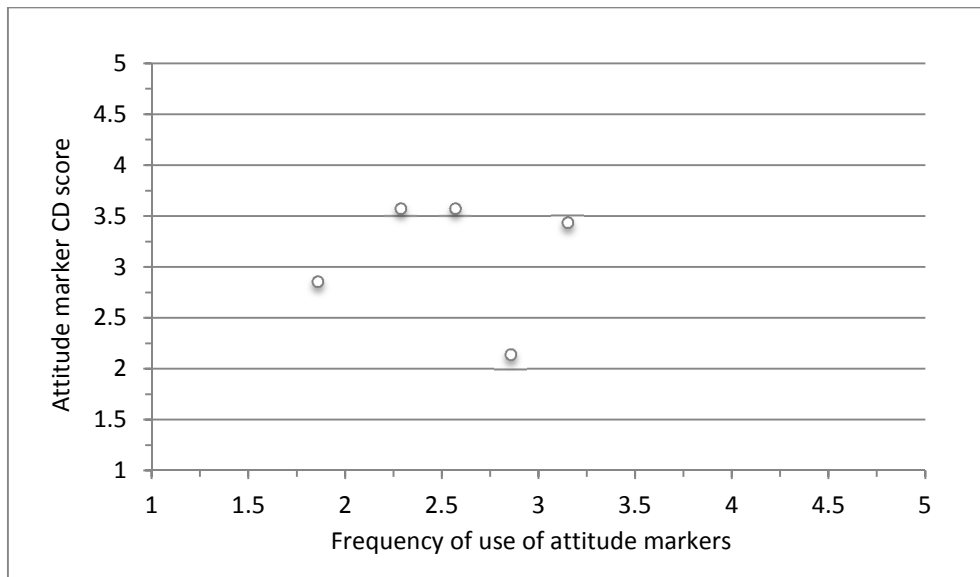
*Figure S1.*

Relationship between the frequency of use of hedges in writing tasks and the concept development (CD) of hedges in writing tasks across participants ( $n = 7$ ).



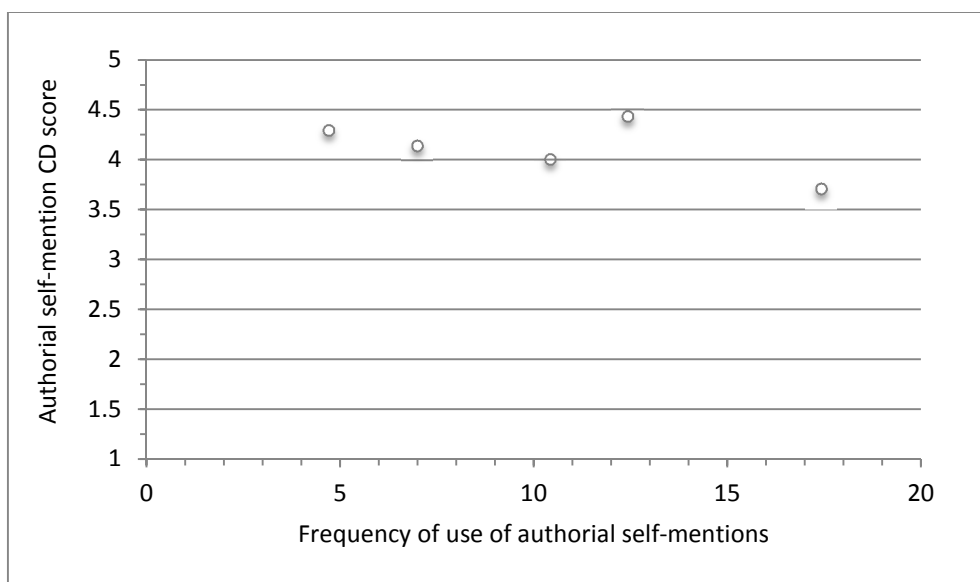
*Figure S2.*

Relationship between the frequency of use of boosters in writing tasks and the concept development (CD) of boosters in writing tasks across participants ( $n = 7$ ).



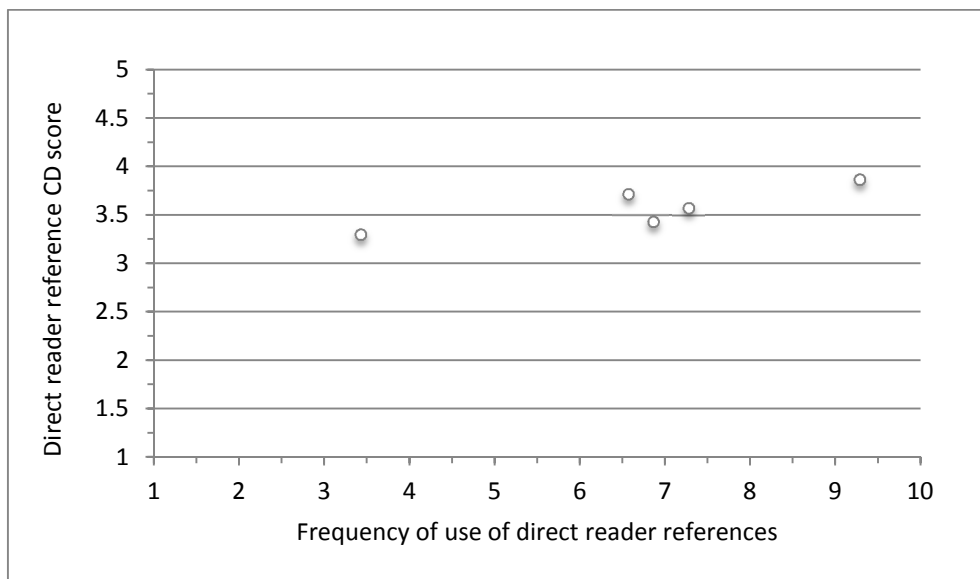
*Figure S3.*

Relationship between the frequency of use of attitude markers in writing tasks and the concept development (CD) of attitude markers in writing tasks across participants ( $n = 7$ ).



*Figure S4.*

Relationship between the frequency of use of authorial self-mention in writing tasks and the concept development (CD) of authorial self-mention in writing tasks across participants ( $n = 7$ ).



*Figure S5.*

Relationship between the frequency of use of direct reader references in writing tasks and the concept development (CD) of direct reader references in writing tasks across participants ( $n = 7$ ).

### Appendix T

Scatterplots for RQ4: Rhetorical Feature Concept Development Scores and Speaking Tasks

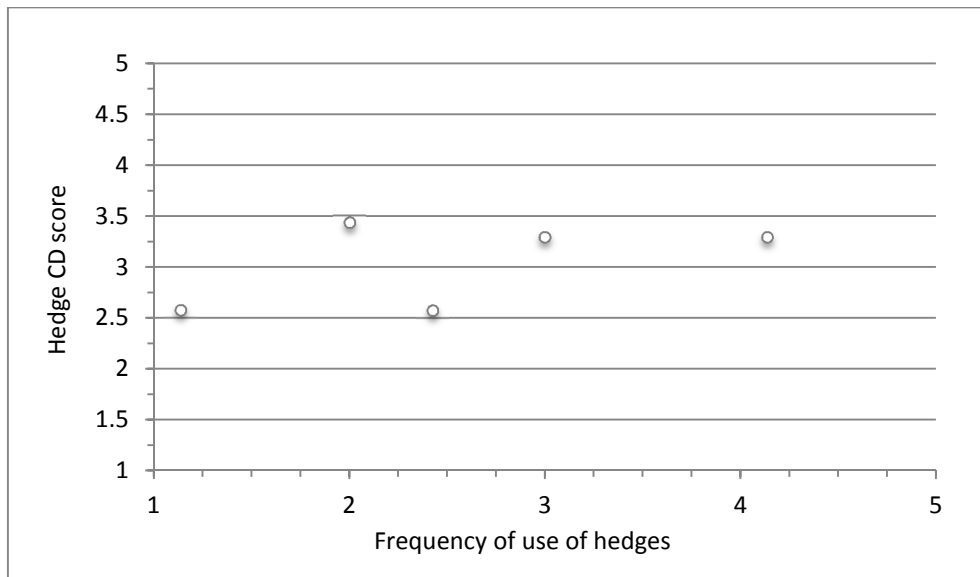


Figure T1.

Relationship between the frequency of use of hedges in writing tasks and the concept development (CD) of hedges in speaking tasks across participants ( $n = 7$ ).

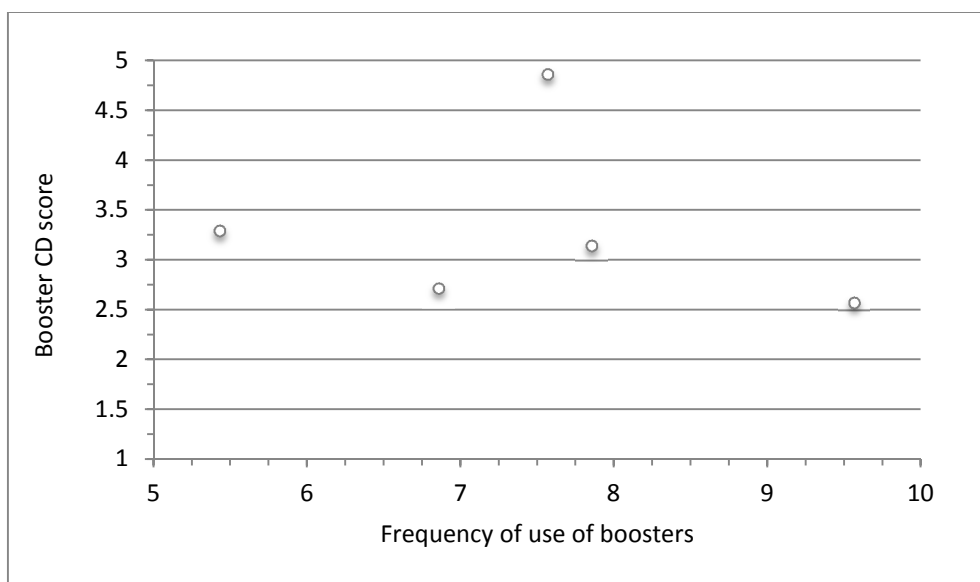


Figure T2.

Relationship between the frequency of use of boosters in writing tasks and the concept development (CD) of boosters in speaking tasks across participants ( $n = 7$ ).

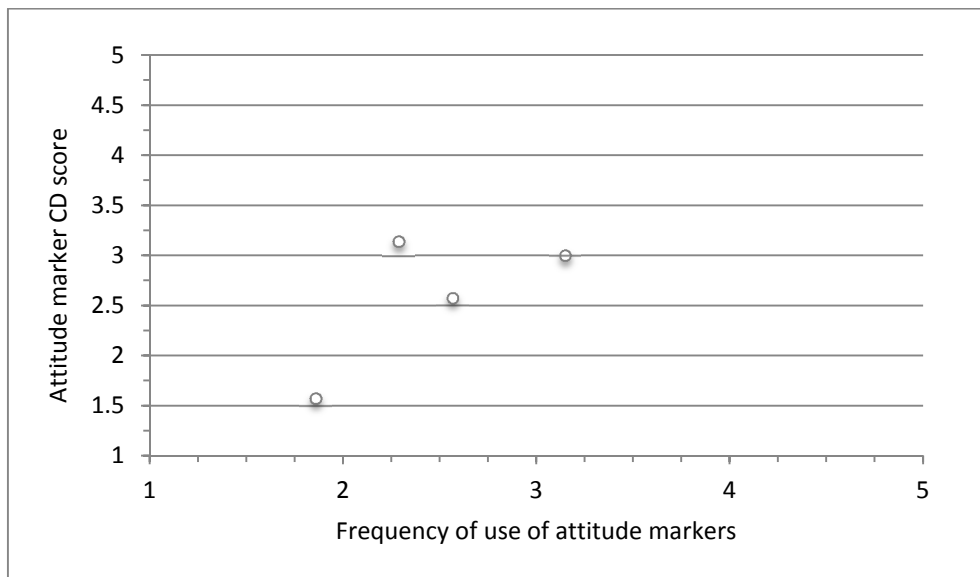


Figure T3.

Relationship between the frequency of use of attitude markers in writing tasks and the concept development (CD) of attitude markers in speaking tasks across participants ( $n = 7$ ).

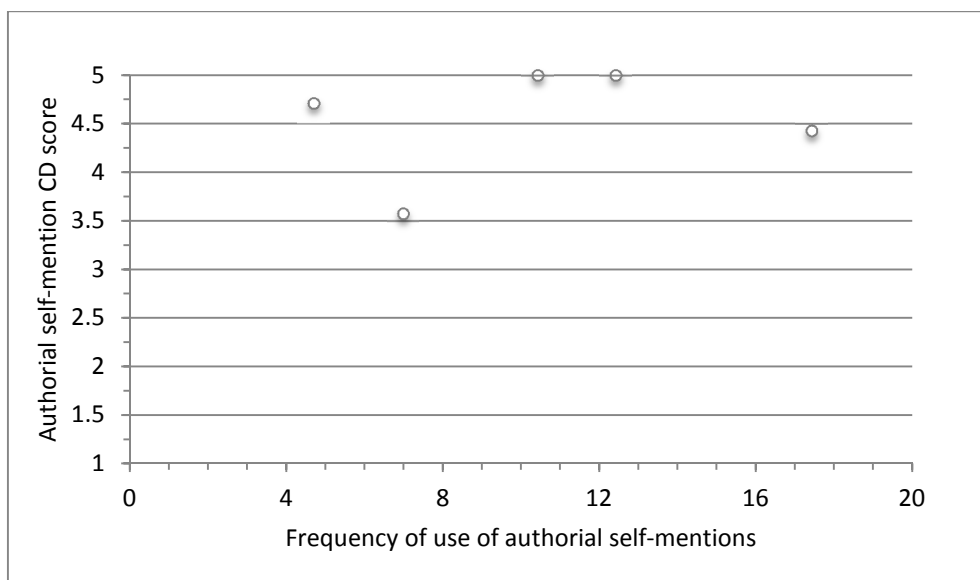
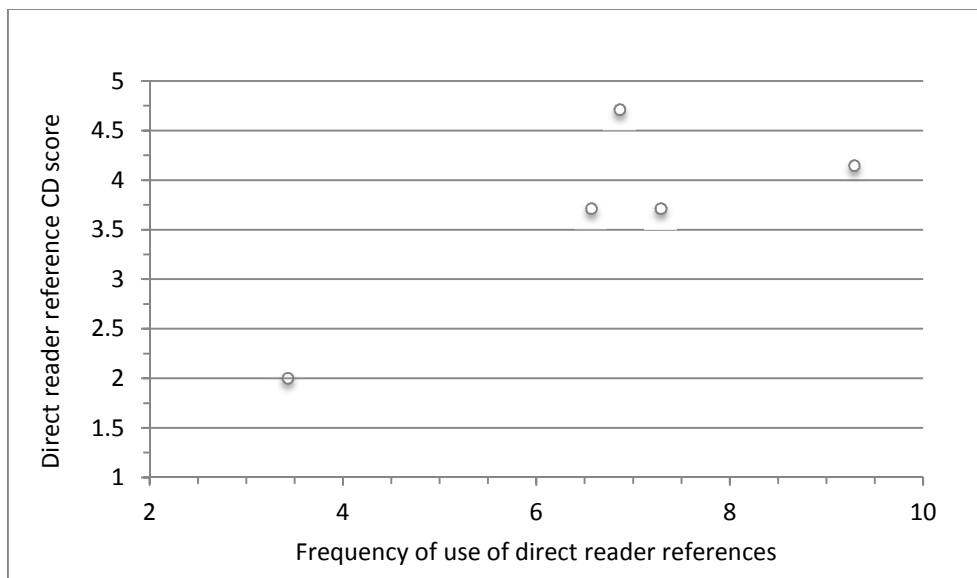


Figure T4.

Relationship between the frequency of use of authorial self-mention in writing tasks and the concept development (CD) of authorial self-mention in speaking tasks across participants ( $n = 7$ ).



*Figure T5.*

Relationship between the frequency of use of direct reader references in writing tasks and the concept development (CD) of direct reader references in speaking tasks across participants ( $n = 7$ ).



## Appendix U Transcription Conventions

Courier New font: direct speech.

Double parentheses: interviewer's comments, such as ((Right.)).

Brackets [ ]: addition of relevant information.

Brackets with italics [*italics*]: citing a writing sample for clear reference (such as sentence from a participant's essay) under discussion.

Ellipsis in brackets [...]: irrelevant information has been left out.

Single inverted commas with italicized script 'text': the speaker is quoting directly from the writing sample under discussion.

Plus sign (+): untimed pause.

Transcription conventions. From *Case study research in applied linguistics* (p. 157), P. A. Duff, 2008, New York: Erlbaum. Copyright 2008 by Erlbaum. Adapted with permission.

## Appendix V

### Summary of Research Findings Organized by Research Question

Research question	Results
<p>1. Is there an increase, over the intervention period, in participants' (a) overall quality of writing, and (b) overall quality of voice in writing?</p>	<ul style="list-style-type: none"> <li>• There are statistically significant increases over the intervention in the overall quality of writing (<math>\chi^2(6) = 20.86, p = .002</math>)—with a large effect size (<math>W = .5</math>)—and the overall quality of voice in writing (<math>\chi^2(6) = 21.37, p = .002</math>)—with a large effect size (<math>W = .51</math>)—across the intervention group.</li> <li>• Post hoc pairwise analyses regarding the overall quality of voice in writing were able to locate the presence of statistically significant increases over time between the following: Day 10 and the posttest, <math>p = .016</math> (<math>p &lt; .025</math>), and the posttest and the delayed posttest, <math>p = .042</math> (<math>p &lt; .05</math>).</li> <li>• Post hoc pairwise analyses were unable to locate the presence of statistically significant increases over time regarding the overall quality of writing.</li> <li>• Effect sizes across both sets of pairwise post hoc analyses (42 comparisons in total) revealed large (in 6 instances) and very large (in 17 instances) effect sizes.</li> <li>• Participants marked stylistics as the primary cause of changes in writing and voice scores. Alongside this lack in variability, this marking indicates that stylistics is the source of the microgenetic changes documented in learners' writing.</li> <li>• All participants were aware of their role as authors in a dialogic space; learners' noted that this awareness was most influential in developing CD<sup>a</sup> of direct reader reference and authorial self-mention.</li> <li>• All participants were aware of the benefits of studying target voice features on their writing.</li> <li>• Participants marked explicit instruction (in stylistic analyses, sample essay analyses, and teacher instructions on effective writing) as effective for improving authorial voice.</li> <li>• The effectiveness of the OBA<sup>b</sup> as a potential source for the statistically significant changes is inconclusive.</li> </ul>
<p>Is any increase in the quality of writing and the quality of voice sustained after the intervention?</p>	<ul style="list-style-type: none"> <li>• Concerning the overall quality of writing, there was no change over time from the posttest to the delayed posttest, <math>z = -0.45, ns (.66)</math>, with a small effect size (<math>r = -.17</math>). This indicates that the increase in the quality of writing observed during the intervention phase was sustained during the four weeks between the posttest and the delayed posttest.</li> <li>• Concerning the overall quality of voice in writing, there was a statistically significance change over time, <math>z = -2.03, p = .04</math>, with a large effect size (<math>r = -.77</math>). This indicates that the increase in the quality of voice in writing was not sustained during the four weeks between the posttest and the delayed posttest.</li> </ul>

2. What is the relationship between the frequency of use of each target rhetorical feature of voice, and (a) the overall quality of writing and (b) the overall quality of voice?

- Regarding voice scores that were not sustained, one participant suggested a longer period of study for internalizing the target features.
- The rate of change concerning voice development was gradual rather than abrupt.
- Regarding the overall quality of writing, there is a positive correlation ( $r_s = .43, p = .34$ ) between the use of hedges and the overall quality of writing. There is a negative correlation ( $r_s = -.64, p = .12$ ) between the use of attitude markers and the overall quality of writing. None of these correlations, however, proved to be statistically significant. Overall, there is no clear pattern of correlation between the number of rhetorical features used and the overall quality of writing, or the type of rhetorical features used and the overall quality of writing
- Regarding the overall quality of voice in writing, there is a positive correlation between the use of boosters ( $r_s = .64, p = .12$ ) and the use of direct reader references ( $r_s = .64, p = .12$ ) with the overall quality of voice in writing. There is a positive correlation regarding the use of authorial self-mention ( $r_s = .57, p = .18$ ). There is a negative correlation regarding the use of attitude markers ( $r_s = -.54, p = .22$ ). Overall, there is a varying degree of positive correlation (barring the use of attitude markers) between rhetorical feature use and the overall quality of voice in writing, although none of these correlations were statistically significant.
- Radar graphs for each participant indicate erratic uses of the rhetorical features across all of the participants, indicating nonlinear cognitive development.
- A graph depicting the frequency of use of each rhetorical feature demonstrates that uses of authorial self-mention, boosters, and direct reader references were more frequent than uses of hedges or attitude markers.

3. What is the relationship between (a) the overall quality of writing and (b) the overall quality of voice in writing?

- There is a positive correlation between the overall quality of writing and the overall quality of voice in writing ( $r_s = .71$ ). However, this correlation was not statistically significant.
- Participants were aware that projecting a strong authorial voice could improve their overall writing; participants also noted a marked effort to engage this element in their writing.
- Participants were aware of the function of all the rhetorical features, with a focus on authorial self-mention and direct reader reference.
- Stylistics is effective as a mediating tool for improving voice across all participants; for Maru, Tom, Anny, Pi, and Am stylistics is also an effective mediating tool for improving overall quality of writing
- Group discussions influenced CD of the target rhetorical features and the overall quality of participant writing.

4. What is the relationship between (a) the frequency of use of each target rhetorical feature of voice and (b) the concept development of each rhetorical feature of voice?

- There is a positive correlation ( $r_s = .67, p = .22$ ) between the use of hedges and the concept development of hedges in writing, and a positive correlation ( $r_s = .7, p = .19$ ) between the use of direct reader references and the concept development of direct reader references in writing. There is also a negative correlation ( $r_s = -.4, p = .51$ ) between the use of authorial self-mention and the concept development of authorial self-mention in writing. These correlations were not statistically significant.
- There is also a positive correlation ( $r_s = .62, p = .27$ ) between the use of direct reader references in writing tasks and the concept development of direct reader references in speaking tasks, and a negative correlation ( $r_s = -.5, p = .39$ ) between the use of boosters in writing tasks and the concept development of boosters in speaking tasks. These correlations were not statistically significant.
- There are more rhetorical features that demonstrate a positive correlation between rhetorical feature use and concept development in writing tasks than in speaking tasks.
- There is a disconnect between what the participants were able to conceptualize and their use of the target features.
- Learners demonstrated an iterative path of cognitive uptake regarding the target discourse features.

5. Are there differences in (a) the quality of writing and (b) the quality of voice in writing between the control group and the course participants?

- There are no differences between the control group and the intervention group regarding the quality of writing and the quality of voice in writing. This result is based on comparisons between written compositions collected as part of the baseline data and the delayed posttest.
- Regarding effect size, comparisons between the quality of writing in the delayed posttest writing samples of the control group and the intervention group indicate a moderate effect size ( $r = -.48$ ).
- Regarding effect size, comparisons between the quality of voice in the delayed posttest writing samples of both the control group and the intervention group indicate a moderate effect size ( $r = -.51$ ).

---

<sup>a</sup> Concept development.

<sup>b</sup> Orienting basis of action.