Writing Development of Second Language Learners of Japanese at an Australian University

Yuka Kikuchi

A thesis submitted for the degree of

Doctor of Philosophy

School of Languages and Linguistics

The University of Melbourne

October 2017
Declaration of Originality

I certify that this thesis does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

Signed: [Signature] On: 23/10/2017
Acknowledgements

The writing and completion of this thesis would not have been possible without the support of many teachers, friends and family. First, I would like to express my deep gratitude to my supervisors, Associate Professor Neomy Storch and Dr Ute Knoch, for their constant support, mentorship and encouragement. I would also like to thank all of my research participants, who generously offered their time to participate in data collection.

I am also grateful to the Japanese Department coordinator of the university, where the data collection was conducted, for being supportive throughout the process. Many thanks go to Christine Eastman for her statistical advice and Alicia Chea, Jenetta Haim, Sara Kitaoji (AE) and Larissa Tittl (AE), who have proofread and edited my work during my PhD. This thesis has been edited in accordance with the ‘Guidelines for editing research theses’, developed jointly by the Institute of Professional Editors (Australia) and Australian universities.

I also want to thank all my friends and colleagues who have been especially supportive throughout my PhD journey. I am thankful to Marzooq Aldossary, Maki Yoshida and Makoto Abe for inspiring me with ideas. Finally and most importantly, I would like to thank my parents for their unconditional support.
Contents

Acknowledgements ........................................................................................................ iii

Contents ........................................................................................................................... iv

List of Figures ................................................................................................................ xi

List of Tables .................................................................................................................. xii

List of Extracts ................................................................................................................ xvi

List of Abbreviations ..................................................................................................... xvii

Abstract .......................................................................................................................... xviii

Chapter 1: Introduction .................................................................................................. 1

1.1 Statement of the Problem ....................................................................................... 3

1.2 Significance of the Study ....................................................................................... 4

1.3 Thesis Chapter Overview ....................................................................................... 6

Chapter 2: Literature Review ......................................................................................... 9

2.1 L2 Writing Development: Theoretical Perspectives .............................................. 9

2.1.1 Dynamic systems theories .............................................................................. 10

2.1.2 Sociocultural theory ..................................................................................... 12

2.1.3 Multi-competence theory .............................................................................. 14

2.1.4 Goal theory .................................................................................................. 15

2.1.5 Genre theory ................................................................................................ 16

2.2 Theories Informing the Current Study—L2 Writing Development from Second Language Acquisition Cognitive Perspectives .................................................. 19

2.2.1 Noticing hypothesis ...................................................................................... 19

2.2.2 Notion of ‘practice’ in skills acquisition theory ............................................ 21

2.3 Approaches to Investigating L2 Writing Development ........................................ 22

2.3.1 Cross-sectional design ................................................................................... 22

2.3.1.1 Cross-sectional studies in an English a second language and English as a Foreign Language context ............................................................... 23

2.3.1.2 Cross-sectional studies of languages other than English ....................... 26

2.3.2 Longitudinal design ...................................................................................... 30

2.3.2.1 Studies on L2 writing development with/without instruction .................. 31

2.4 Students’ Perceptions of Writing Development ................................................... 39

2.5 Factors that Affect Writing Development ............................................................. 43

2.5.1 Instruction ..................................................................................................... 43

2.5.1.1 Product approach ................................................................................... 43

2.5.1.2 Process approach ................................................................................... 45

2.5.1.3 Genre approach ................................................................................... 46

2.5.1.4 Process-genre approach ....................................................................... 48

2.5.2 Feedback ....................................................................................................... 49

2.5.2.1 Teacher feedback ................................................................................... 49

2.5.2.2 Students’ beliefs about teacher feedback .............................................. 53

2.5.2.3 Peer feedback ....................................................................................... 54

2.5.2.4 Factors affecting the benefits of peer feedback ..................................... 56

2.5.3 Reading and writing ....................................................................................... 58

2.5.3.1 Reading and writing connections ............................................................ 59

2.5.3.2 Vocabulary acquisition and extensive reading .................................... 63

2.5.4 Motivation and L2 writing ............................................................................. 64

2.5.5 Language background ................................................................................... 68
2.5.1 Background speaker and non-background speaker learners of languages other than Japanese ........................................... 69
2.5.2 Background speaker and non-background speaker learners of Japanese .............................................................. 72
2.6 Gaps in the Research ........................................................................................................................................ 75
2.7 The Present Study ........................................................................................................................................ 77

Chapter 3: Discourse Analytic Measures ........................................................................................................ 79
3.1 Features of the Japanese Language ........................................................................................................ 79
3.2 Complexity, Accuracy and Fluency ........................................................................................................ 83
  3.2.1 Complexity ........................................................................................................................................ 84
    3.2.1.1 Lexical complexity ........................................................................................................ 84
    3.2.1.2 Syntactic complexity ...................................................................................................... 89
  3.2.2 Accuracy ........................................................................................................................................ 97
  3.2.3 Fluency ........................................................................................................................................ 102
    3.2.3.1 Measures of temporal aspects of fluency .................................................................. 103
3.3 Other Measures that Examine Writing Development ........................................................................ 107
  3.3.1 Mechanics ........................................................................................................................................ 107
  3.3.2 Cohesion and coherence ................................................................................................................. 107
    3.3.2.1 Cohesion .......................................................................................................................... 108
    3.3.2.2 Coherence .......................................................................................................................... 110
  3.3.3 Desu/masu form and plain form .................................................................................................... 113
3.4 The Chosen Discourse Analytic Measures ........................................................................................... 113

Chapter 4: Study 1 Methodology ................................................................................................................ 115
4.1 Study 1 Design ........................................................................................................................................ 115
4.2 Participants ............................................................................................................................................. 116
4.3 Instruments ............................................................................................................................................ 117
  4.3.1 Placement test/writing task ........................................................................................................ 117
  4.3.2 Background questionnaire ............................................................................................................ 117
4.4 Procedures ............................................................................................................................................ 118
  4.4.1 Pilot study ...................................................................................................................................... 118
    4.4.1.1 Data-coding procedure and pilot study results ..................................................... 119
  4.4.2 Statistical analysis ......................................................................................................................... 133
  4.4.3 Inter-coder reliability analysis for Study 1 ................................................................................ 134

Chapter 5: Study 1 Results and Discussion .................................................................................................. 136
5.1 Statistical Analysis: Correlations ........................................................................................................ 136
5.2 Results and Discussion ........................................................................................................................ 138
  5.2.1 Kanji complexity .......................................................................................................................... 139
    5.2.1.1 Kanji character root type token ratio ........................................................................... 139
    5.2.1.2 Kanji character corrected type token ratio ............................................................. 140
    5.2.1.3 Kanji word root type token ratio .............................................................................. 142
    5.2.1.4 Kanji word corrected type token ratio ................................................................. 143
  5.2.2 Kanji complexity results summary ............................................................................................. 144
  5.2.3 Syntactic complexity ....................................................................................................................... 145
    5.2.3.1 Ratio of te-form to words .......................................................................................... 145
    5.2.3.2 Ratio of complex sentences to sentences ................................................................. 147
    5.2.3.3 Ratio of clause conjunctions to words ....................................................................... 149
    5.2.3.4 Ratio of propositional endings to words ................................................................. 150
    5.2.3.5 Ratio of epistemic endings to words ..................................................................... 152
    5.2.3.6 Ratio of expressive endings to words .................................................................... 154
  5.2.4 Syntactic complexity results summary ....................................................................................... 156
  5.2.5 Accuracy ....................................................................................................................................... 157
Chapter 6: Study 2 Methodology

6.1 Context ............................................. 172
  6.1.1 Japanese 3 (Semester 1) .................................. 172
  6.1.2 Japanese 4 (Semester 2) .................................. 174
6.2 Study Design ........................................... 175
  6.2.1 Phase 1 design: one semester—Japanese 3 .................. 175
  6.2.2 Phase 2 design: one year (Japanese 3 & Japanese 4) ........ 177
  6.2.3 Participants .......................................... 178
    6.2.3.1 Student participants .................................. 178
    6.2.3.2 Teacher participants .................................. 179
6.3 Data Collection Tools ..................................... 179
  6.3.1 Pre- and post- writing tests ................................ 180
  6.3.2 Semi-structured interviews .................................. 180
  6.3.3 Background questionnaire .................................... 182
  6.3.4 Lecture and classroom observation .............................. 182
  6.3.5 Individual assignments ....................................... 183
6.4 Data-Coding Procedures ................................... 183
  6.4.1 Content scale ........................................... 185
    6.4.1.1 Sufficiency ............................................ 186
    6.4.1.2 Persuasiveness .......................................... 186
  6.4.2 Structure .............................................. 188
    6.4.2.1 Coherence ............................................. 188
    6.4.2.2 Paragraphing ............................................ 188
6.5 Reliability Analysis ....................................... 189
6.6 Data Analysis Procedure ................................... 190
  6.6.1 Discourse analytic measures and rating scales analysis ........ 190
  6.6.2 Analysing interviews ....................................... 191
    6.6.2.1 Analysis of student interviews .......................... 192
    6.6.2.2 Analysis of teacher interviews .......................... 193

Chapter 7: Study 2 Text-Based Analyses Results .................................. 194
7.1 Kanji Complexity: Short- and Long-term Results .................. 194
7.2 Syntactic Complexity: Short- and Long-term Results .............. 196
7.3 Accuracy: Short- and Long-term Results ................................ 198
7.4 Fluency: Short- and Long-term Results ................................ 199
7.5 Content: Short- and Long-term Results ................................ 201
7.6 Structure: Short- and Long-term Results ............................. 202
7.7 Summary of the Main Findings Related to Research Question 1 .... 204
7.8 Character-based Learners and Non-character-based Learners ....... 205
    7.8.1 Kanji complexity ......................................... 206
    7.8.2 Syntactic complexity ...................................... 209
10.3 Limitations of the Study .................................................................330
10.4 Suggestion for Further Research ..................................................331
10.5 Concluding Statement .................................................................332

References .......................................................................................334

Appendices .......................................................................................369
List of Figures

Figure 3.1. Example of the word ‘Kyoto’ (京都) .......................................................... 106
Figure 5.1. Distribution of kanji character root type token ratio per writing script over five different levels ................................................................. 140
Figure 5.2. Distribution of kanji character corrected type token ratio per writing script over five different levels ................................................................. 141
Figure 5.3. Distribution of kanji word root type token ratio per writing script over five different levels ................................................................. 142
Figure 5.4. Distribution of kanji word corrected type token ratio per writing script over five different levels ................................................................. 144
Figure 5.5. Distribution of the ratio of te-form to words per writing script over five different levels ................................................................. 146
Figure 5.6. Distribution of ratio of complex sentences to sentences per writing script over five different levels ................................................................. 148
Figure 5.7. Distribution of ratio of clause conjunctions per writing script over five different levels ................................................................. 149
Figure 5.8. Distribution of ratio of propositional endings to words per writing script over five different levels ................................................................. 151
Figure 5.9. Distribution of ratio of epistemic endings to words per writing script over five different levels ................................................................. 153
Figure 5.10. Distribution of ratio of expressive endings to words per writing script over five different levels ................................................................. 155
Figure 5.11. Distribution of ratio of error-free sentences per writing script over five different levels ................................................................. 158
Figure 5.12. Distribution of ratio of error-free clauses per writing script over five different levels ................................................................. 160
Figure 5.13. Distribution of words per error-free sentences per writing script over five different levels ................................................................. 161
Figure 5.14. Distribution of number of words per writing script over five different levels ................................................................. 163
Figure 5.15. Distribution of number of sentences per writing script over five different levels ................................................................. 165
Figure 5.16. Distribution of number of kanji characters per writing script over five different levels ................................................................. 166
Figure 5.17. Distribution of number of kanji words per writing script over five different levels ................................................................. 168
Figure 5.18. Distribution of number of clauses per writing script over five different levels ................................................................. 169
Figure 7.1. Estimated marginal means for kanji complexity .............................................. 207
Figure 7.2. Estimated marginal means for syntactic complexity .............................................. 210
Figure 7.3. Estimated marginal means for accuracy ............................................................ 212
Figure 7.4. Estimated marginal means for fluency ............................................................ 214
Figure 7.5. Estimated marginal means for content. ............................................. 216
Figure 7.6. Estimated marginal means for structure ........................................ 218
Figure 9.1. Slide on explanation of coherence. .................................................... 306
List of Tables

Table 2.1 Studies on Transferability of Text Features across L1 and L2 ........................................14
Table 2.2 Comparison of Results on Studies Using Statistical Significance: Treatment Group Gains Over Control Groups Gains .................................................................61
Table 3.1 Example of Kanji 本 .........................................................................................80
Table 3.2 Summary of the Japanese Writing System .................................................................81
Table 3.3 Breakdown of Content Words and Grammatical Words ........................................82
Table 3.4 Summary of Japanese Words ..................................................................................83
Table 3.5 Complexity Measures Examined in Iwashita’s (2006) Study ................................91
Table 3.6 Eight Types of Sentence-ending Variables ...............................................................96
Table 3.7 Types of Kanji Errors .........................................................................................100
Table 3.8 Types of Kana Errors .........................................................................................101
Table 3.9 Summary of Measures .......................................................................................114
Table 4.1 Distribution of Test Papers from the Different Levels ........................................116
Table 4.2 First Languages of Participants ............................................................................116
Table 4.3 Kanji Complexity .................................................................................................119
Table 4.4 Results for Various Kanji Complexity Measures ................................................121
Table 4.5 Syntactic Complexity Measures ........................................................................121
Table 4.6 Results for the Syntactic Complexity Measures ................................................124
Table 4.7 Sentence-ending Variables and Examples .............................................................125
Table 4.8 Results for Various Sentence-ending Variables ................................................126
Table 4.9 Accuracy Measures .............................................................................................127
Table 4.10 Results for Various Accuracy Measures ..........................................................129
Table 4.11 Fluency Measures ..............................................................................................130
Table 4.12 Results for Various Fluency Measures ..............................................................131
Table 4.13 Cohesion Results ...............................................................................................132
Table 4.14 Measures Used in Study 1 ................................................................................133
Table 4.15 Kanji Complexity Inter-rater Reliability.........................................................134
Table 4.16 Syntactic Complexity Inter-rater Reliability......................................................135
Table 4.17 Accuracy Inter-rater Reliability .......................................................................135
Table 4.18 Fluency Inter-rater Reliability .........................................................................135
Table 4.19 Cohesion Inter-rater Reliability .......................................................................135
Table 5.1 Results of Bivariate Pearson Correlation Tests.....................................................137
Table 5.2 Descriptive Statistics—Kanji Complexity: Kanji Character Root Type Token Ratio .........................................................................................................................139
Table 5.3 Descriptive Statistics—Kanji Complexity: Kanji Character Corrected Type Token Ratio ..........................................................................................................................141
Table 5.4 Descriptive Statistics: Kanji Complexity—Kanji Word Root Type Token Ratio ........................................................................................................................................142
Table 5.5 Descriptive Statistics—Kanji Complexity: Kanji Word Corrected Type Token Ratio ................................................................................................................................143
Table 5.6 Descriptive Statistics—Syntactic Complexity: Ratio of Te-form to Words ..........146
Table 5.7 Descriptive Statistics—Syntactic Complexity: Ratio of Complex Sentences ......147
Table 5.8 Descriptive Statistics—Syntactic Complexity: Ratio of Clause Conjunctions to Words ...............................................................................................................................149
Table 5.9 Descriptive Statistics—Syntactic Complexity: Ratio of Propositional Endings ...151
Table 5.10 Descriptive Statistics—Syntactic Complexity: Ratio of Epistemic Endings......153
Table 5.11 Descriptive Statistics—Syntactic Complexity: Ratio of Expressive Endings .....154
Table 5.12 Descriptive Statistics—Accuracy: Ratio of Error-free Sentences ......................158
Table 5.13 Descriptive Statistics—Accuracy: Ratio of Error-free Clauses .........................159
Table 5.14 Descriptive Statistics—Accuracy: Words per Error-free Sentences...............161
Table 5.15 Descriptive Statistics—Average Number of Words per Script .........................163
Table 5.16 Descriptive Statistics—Average Number of Sentences per Script..................164
Table 5.17 Descriptive Statistics—Average Number of Kanji Characters per Script .........166
Table 5.18 Descriptive Statistics—Average Number of Kanji Words per Script ...............167
Table 5.19 Descriptive Statistics—Average Number of Clauses per Script..........................169
Table 6.1 Japanese 3 Assessments and Requirements .......................................................173
Table 6.2 Japanese 4 Assessments and Requirements .......................................................174
Table 6.3 Schedule for Japanese 3 Data Collection ..........................................................176
Table 6.4 Schedule for Japanese 4 Data Collection ..........................................................177
Table 6.5 Character-based and Non-character-based Learners in Japanese 3 and
Japanese 4 .........................................................................................................................178
Table 6.6 Discourse Analytic Measures .............................................................................184
Table 6.7 Content Scale ....................................................................................................187
Table 6.8 Coherence/Logical Development .....................................................................189
Table 6.9 Paragraphing .....................................................................................................189
Table 6.10 Broad Themes and Sub-themes Identified for Student Interviews .................192
Table 6.11 Broad Themes and Sub-themes Identified for Teacher Interviews ...............193
Table 7.1 Descriptive Statistics for the Short and Long Term: Kanji Complexity.............195
Table 7.3 Descriptive Statistics for Short and Long Term: Syntactic Complexity .............197
Table 7.4 Repeated Measure ANOVA Results for Short and Long Term: Syntactic
Complexity .........................................................................................................................198
Table 7.5 Descriptive Statistics for Short and Long Term: Accuracy ...............................198
Table 7.6 Repeated Measures ANOVA Results for Short and Long Term: Accuracy .......199
Table 7.7 Descriptive Statistics for the Short and Long Term: Fluency .............................200
Table 7.8 Repeated Measures ANOVA Results for Short and Long Term: Fluency ........201
Table 7.9 Descriptive Statistics for Short and Long Term: Content (Sufficiency and
Persuasiveness) .................................................................................................................202
Table 7.10 Repeated Measures ANOVA Results for the Short and Long Term: Content ...202
Table 7.11 Descriptive Statistics for the Short and Long Term: Structure (Coherence
and Paragraphing) .........................................................................................................203
Table 7.12 Repeated Measures ANOVA Results for the Short and Long Term: Structure ..203
Table 7.13 Descriptive Statistics for Character-based and Non-character-based Learners:
Kanji Complexity

Table 7.14 Repeated Measures ANOVA Results for Character-based and Non-character-based Learners: Kanji Complexity

Table 7.15 Pairwise Comparison: Mean difference for Kanji Complexity

Table 7.16 Descriptive Statistics for Character-based and Non-character-based Learners: Syntactic Complexity

Table 7.18 Descriptive Statistics for Character-based and Non-character-based Learners: Accuracy

Table 7.19 Repeated Measures ANOVA Results for Character-based and Non-character-based Learners: Accuracy

Table 7.20 Descriptive Statistics for Character-based and Non-character-based Learners: Fluency

Table 7.21 Repeated Measures ANOVA Results for Character-based and Non-character-based Learners: Fluency

Table 7.22 Pairwise Comparison: Mean Difference for No. of KCs and KWs

Table 7.23 Descriptive Statistics for Character-based and Non-character-based Learners: Content

Table 7.24 Repeated Measures ANOVA Results for Character-based and Non-character-based Learners: Content

Table 7.25 Descriptive Statistics for Character-based and Non-character-based Learners: Structure

Table 7.26 Repeated Measures ANOVA Results for Character-based and Non-character-based Learners: Structure

Table 8.1 Factors that Influenced Students’ Writing Development

Table 8.2 Peer Feedback Criteria

Table 9.1 Short-term Results (One Semester)

Table 9.2 Long-term Results (One Year)

Table 9.3 Summary of Developmental Changes in Study 1 and Study 2
List of Extracts

Extract 9.1 One Section of Jasmine’s Japanese 3 Pre-test Writing Script ............... 296
Extract 9.2 One Section of Jasmine’s Japanese 3 Post-test Writing Script ............. 296
Extract 9.3 One Section of Mary’s Individual Assignment .................................. 297
Extract 9.4 One Section of Ricki’s Individual Assignment ................................. 298
Extract 9.5 One Section of David’s Japanese 3 Post-test Writing Script ............... 298
Extract 9.6 One Section of Sharon’s Japanese 3 Post-test Writing Script .......... 299
Extract 9.7 One Section of Helena’s Japanese 3 Pre-test Writing Script ............ 301
Extract 9.8 One Section of Helena’s Japanese 3 Post-test Writing Script ............ 301
Extract 9.9 One Section of Peter’s Japanese 3 Pre-test Writing Script ............... 304
Extract 9.10 One Section of Peter’s Japanese 3 Post-test Writing Script .......... 304
Extract 9.11 One Section of Peter’s Japanese 4 Post-test Writing Script ............. 305
List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTFL</td>
<td>American Council on the Teaching of Foreign Languages</td>
</tr>
<tr>
<td>ACT-R</td>
<td>Adaptive control of thought-rational theory</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of variance</td>
</tr>
<tr>
<td>AWL</td>
<td>Academic word list</td>
</tr>
<tr>
<td>BS</td>
<td>Background speaker</td>
</tr>
<tr>
<td>CAF</td>
<td>Complexity, accuracy and fluency</td>
</tr>
<tr>
<td>CTTR</td>
<td>Corrected type token ratio</td>
</tr>
<tr>
<td>DELA</td>
<td>Diagnostic English Language Assessment</td>
</tr>
<tr>
<td>DMM</td>
<td>Dynamic model of multilingualism</td>
</tr>
<tr>
<td>DST</td>
<td>Dynamic systems theories</td>
</tr>
<tr>
<td>DUB</td>
<td>Dynamic usage-based theory</td>
</tr>
<tr>
<td>EAP</td>
<td>English for academic purposes</td>
</tr>
<tr>
<td>EFC</td>
<td>Error-free clause</td>
</tr>
<tr>
<td>EFL</td>
<td>English as a foreign language</td>
</tr>
<tr>
<td>EFT</td>
<td>Error-free T-unit</td>
</tr>
<tr>
<td>ELP</td>
<td>English-language proficiency</td>
</tr>
<tr>
<td>EPP</td>
<td>Extended parallel progression</td>
</tr>
<tr>
<td>ER</td>
<td>Extensive reading</td>
</tr>
<tr>
<td>ESL</td>
<td>English as a second language</td>
</tr>
<tr>
<td>EYL</td>
<td>English for Young Learners</td>
</tr>
<tr>
<td>FL</td>
<td>Foreign language</td>
</tr>
<tr>
<td>IELTS</td>
<td>International English Language Testing System</td>
</tr>
<tr>
<td>JFL</td>
<td>Japanese as a foreign language</td>
</tr>
<tr>
<td>JSL</td>
<td>Japanese as a second language</td>
</tr>
<tr>
<td>KC</td>
<td><em>Kanji</em> character</td>
</tr>
<tr>
<td>KW</td>
<td><em>Kanji</em> word</td>
</tr>
<tr>
<td>LOTE</td>
<td>Languages other than English</td>
</tr>
<tr>
<td>LSD</td>
<td>Least significant difference</td>
</tr>
<tr>
<td>NBS</td>
<td>Non-background speaker</td>
</tr>
<tr>
<td>PAD</td>
<td>Presenting academic discourse</td>
</tr>
<tr>
<td>PP</td>
<td>Parallel progression</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>RQ</td>
<td>Research questions</td>
</tr>
<tr>
<td>RTTR</td>
<td>Root type token ratio</td>
</tr>
<tr>
<td>SAALE</td>
<td>Student Achievement in Asian Languages Education</td>
</tr>
<tr>
<td>SFL</td>
<td>Systemic functional linguistics</td>
</tr>
<tr>
<td>SLA</td>
<td>Second language acquisition</td>
</tr>
<tr>
<td>SP</td>
<td>Sequential progression</td>
</tr>
<tr>
<td>TDO</td>
<td>Tourist destination overviews</td>
</tr>
<tr>
<td>TL</td>
<td>Target language</td>
</tr>
<tr>
<td>TOEFL</td>
<td>Test of English as a Foreign Language</td>
</tr>
<tr>
<td>TOEFL iBT</td>
<td>TOEFL internet-based test</td>
</tr>
<tr>
<td>TSA</td>
<td>Topical structural analysis</td>
</tr>
<tr>
<td>TTR</td>
<td>Type token ratio</td>
</tr>
<tr>
<td>TWE</td>
<td>Test of Written English</td>
</tr>
<tr>
<td>US</td>
<td>United States of America</td>
</tr>
<tr>
<td>VCE</td>
<td>Victorian Certificate of Education</td>
</tr>
<tr>
<td>ZPD</td>
<td>Zone of proximal development</td>
</tr>
</tbody>
</table>
Abstract

Although interest in L2 writing development has increased in recent years, researchers who have investigated this area (e.g., Knoch, Macqueen & O’Hagan, 2014; Storch 2009) have done so mostly in an English as a second language (ESL) or English as a foreign language (EFL) context. To date, very little research has been conducted on L2 Japanese writing development. This study, a two-part research project (combining a cross-sectional design with a longitudinal classroom-based mixed-methods design), examines the writing development of university students learning Japanese in Australia.

The first part of this study comprises a cross-sectional design. The purpose here was to determine how the writing abilities of L2 Japanese learners at university vary across different levels of proficiency. The writing scripts used in this study originated from students who had participated in the trial of a new placement test (Semester 1, 2013) in the Japanese Department of a large metropolitan Australian university. One-hundred-and-forty-six written scripts were used for the analysis. Four different writing features were examined: kanji complexity, syntactic complexity, accuracy, and fluency, with each containing a number of measures. The study found that overall, the measures examined for kanji complexity and fluency exhibited a greater number of developmental changes compared to syntactic complexity and accuracy.

The study’s second part consisted of a longitudinal classroom-based study that investigated the writing development of lower-intermediate L2 Japanese tertiary learners (from the same university as the first part of the study). This was further divided into two phases. The first phase (conducted over one semester) involved 41 participants and the second phase (conducted over one year) involved 22 participants. The study investigated: 1) whether the students’ writing developed in the short term (one semester) as well as in the long term (one year), what aspects of writing were
developed (if any) and what factors contributed to this development; 2) whether any
difference was apparent in L2 writing development between character-based learners
and non-character-based learners. The data included pre- and post-tests, interviews
(with students and teachers), background questionnaires and classroom observation
notes.

A comparison of students’ performance on the pre- and post-tests showed that overall,
the students’ writing developed in both the short and long term in terms of kanji
complexity, syntactic complexity, fluency, accuracy, content and structure. However, a
few measures of syntactic complexity and accuracy showed no change in both the short
and long term, or only showed improvement in the short term. When comparing the
writing development of character-based learners and non-character-based learners, the
former group improved more than the latter group in the development of kanji-related
areas, such as in number of kanji characters and words (fluency).

The main findings in the longitudinal study were supported by a qualitative analysis of
the interviews, as well as classroom observations. Several factors led to learners’
writing development, or the lack thereof: practice, reading, and explicit instruction and
feedback (teacher and peer feedback). These findings are informed by two second
language acquisition cognitive theories (the noticing hypothesis [Schmidt, 1990], and
‘practice’ in skills acquisition theory [DeKeyser, 2007]). The findings suggest that
awareness (i.e., noticing) and written output practice is conducive to second language
writing development and acquisition when teachers give instructions that require
students to be aware of certain forms (e.g., a grammatical pattern) and to practice them.
Chapter 1: Introduction

A growing worldwide interest in learning Japanese has recently become apparent, with as many as four million people studying the language globally (Japan Times, 2013). The number of learners studying Japanese worldwide has increased 9.1 per cent in 2012 from 2009 (Japan Times, 2013). Japanese has been a popular language in Australia since the 1970s and is now one of the most widely studied languages in both schools and universities (De Kretser & Spence-Brown, 2010). This interest may stem from the fact that Australia has a close relationship with Japan, with the two countries operating as economic partners, approximately since the 1950s (De Kretser & Spence-Brown, 2010). In addition, the growing interest in Japanese animation, culture and music among young people has contributed to interest in the language (Ministry of Foreign Affairs of Japan, 2009).

Students who are learning Japanese in Australia come from different backgrounds, a result of Australia’s significant multiculturalism. In addition, the number of international students coming to Australian universities has increased exponentially in the past decade. In 2014, international students comprised a substantial proportion (approximately 20%) of the 1.4 million higher education students in Australia (Australian Government Department of Education, 2015a, 2015b). These students come from a wide range of countries; however, a significant proportion originate from Asian countries (Australian Government Department of Education, 2015b).

Language consists of four main elements; writing, reading, speaking and listening. A significant question regards whether writing is important for second language acquisition. Many language teachers prioritise speaking over writing; however, languages that contain character-based scripts are thought to be an exception to the importance of speaking; in these scripts, teaching writing at an early stage is
believed equally important (Cook & Bassetti, 2005). Some scholars (e.g., Marriott, Neustupny & Spence-Brown, 1993) argue that acquisition may take more time for character-based languages, such as Japanese and Chinese, compared to alphabet-based languages. The Japanese writing system contains a combination of kanji (Chinese characters), hiragana and katakana and is very complicated (hiragana and katakana are for the Japanese syllabary and katakana is used for loan words). The Japanese writing system, especially kanji, appears difficult for many second language (L2) learners to conquer (Aoki, 2010; Hatasa, 1989; Iwashita & Sekiguchi, 2009; Kirwan 2005; Mori & Shimizu, 2007). Thus, for learners of Japanese to develop their writing skills in that language, it is important to practice writing at an early stage.

Having taught Japanese as a second language at various proficiency levels, I have encountered many who find writing in Japanese challenging. Most students learn hiragana and katakana, with only some kanji at the beginner stage. However, lower-intermediate L2 Japanese learners must begin accumulating kanji, as well as retaining all the hiragana and katakana they have learned. It is at this stage that they find learning difficult. These lower level intermediate learners should also attain the ability to write a reasonable amount of text in Japanese. In this context, my teaching experience explains my interest and desire to investigate what leads to improvements in Japanese writing, especially at the lower-intermediate level. Moreover, most research on L2 writing development to date has been undertaken in relation to ESL or EFL. Additionally, research on Japanese L2 writing development is relatively scarce. This too has motivated me to investigate Japanese L2 writing development.

Teaching students from different backgrounds encouraged me to investigate whether differences existed in the writing development between character-based and non-character-based learners. The potential of background learners having an advantage over non-background learners (especially in writing) is an important issue, particularly
in a country like Australia, which is not only multicultural, but is also experiencing an increase in the number of international students attending tertiary institutions.

This research aims to examine the writing development of L2 tertiary Japanese learners in Australia. The research has been conducted in two phases, investigating writing development through a cross-sectional and a longitudinal study. The cross-sectional phase (Study 1) investigated writing development across a broad range of proficiency levels. The longitudinal phase (Study 2) examined the writing development of lower level intermediate tertiary Japanese learners in the short (over one semester) and long term (over one year). It also examined the differences in writing development between learners with character-based language backgrounds and learners with non-character-based language backgrounds.

1.1 Statement of the Problem

The study of L2 writing development has experienced considerable growth since the early 1990s (Manchón, 2012). To gain a better understanding of L2 writing developmental stages, researchers (e.g., Banerjee, Franceschina & Smith, 2007; Knoch, Macqueen & O’Hagan, 2014) have exhibited a growing interest in the relationship between writing performance and L2 proficiency. Isolating the different variables and examining learners’ writing performance may explain the differences in learners’ L2 developmental trajectory. Until now, previous studies on L2 writing development have reported on certain measures of students’ writing development over time (e.g., complexity, accuracy and fluency). However, most of these studies focus on ESL/EFL learners (e.g., Knoch, Rouhshad & Storch, 2014; Sasaki 2007, 2009, 2011; Storch 2009; Verspoor, Schmid & Xu, 2012). These studies have tended to adopt a test-retest design using a range of measures to assess writing.

As most L2 writing development studies have focused on ESL/EFL, I was interested in determining what features develop in Japanese L2 writing. Would the areas
of improvement differ depending on the learner’s proficiency level? Do some features (e.g., accuracy, complexity) take more time to develop in Japanese L2 writing? What kind of assessment tasks and classroom activities improve a learner’s writing?

A small body of literature has examined L2 Japanese writing (e.g., Aoki, 2010; Iwashita & Sekiguchi, 2009; Machida, 2001; Pennington & So, 1993, Tanaka, 1997). Some studies investigated why background learners perform better than do non-background learners; however, much of this research (e.g., Aoki, 2010; Machida, 2001) is not focused on writing development, but rather writing performance at one point in time. Some scholars, such as Tanaka (1997), have attempted to investigate the writing development of L2 Japanese learners by examining writing development through in-country experience in Japan. Other scholars, such as Iwashita and Sekiguchi (2009), have examined whether a learner’s background and mode of instruction (whether they had studied the language at school or in an intensive first-year course at university) influenced their writing development. To date, there are very few studies on L2 Japanese writing development and none of these has combined cross-sectional and longitudinal designs.

### 1.2 Significance of the Study

The current project aims to expand our understanding of Japanese L2 writing development. It is hoped that the findings will provide an insight into the writing development of L2 Japanese learners at the tertiary level, which will be useful for Japanese teachers. First, it will provide a better understanding of L2 Japanese writing development, as it has examined writing development by isolating features (e.g., complexity, fluency) and seeing how they differ in learners’ L2 developmental trajectory. Second, some measures examined in this project are specific to the Japanese language; thus, the findings will enable effective guidance about L2 Japanese writing development for teachers of the language.
In order to achieve this aim, two studies (Study 1 and Study 2) were conducted. As mentioned above, Study 1 used a cross-sectional design and Study 2 used a longitudinal design. Both studies were conducted at the same institution, a large metropolitan university in Australia with a large number of students studying Japanese as a second language.

Study 1 explored the writing ability of L2 Japanese learners at the university across different proficiency levels. Five different levels of writing scripts from students who participated in a Japanese placement test were examined to determine the differences in learners’ developmental trajectory. Study 2 examined the writing development of lower level intermediate Japanese learners in an authentic classroom. Moreover, through using classroom observation and interview data, Study 2 also considered possible explanations of development, or the lack thereof, in L2 writing over time.

Study 2 also set out to examine the difference in L2 writing development between character-based learners and non-character-based learners. Character-based learners may have a potential advantage over non-character-based learners when it comes to writing. Therefore, understanding the differences in development between the two groups is important, especially in countries like Australia that have many students from different backgrounds.

Study 2 was informed by two psycholinguistic theories of second language acquisition (SLA): the noticing hypothesis (Schmidt, 1990) and the notion of ‘practice’ in skill acquisition theory (DeKeyser, 1997). Although SLA theories have mainly been used in the past to explain oral skills, in recent years many L2 writing scholars (e.g., Bitchener & Storch, 2016, Williams, 2012) have suggested that SLA theories might inform L2 writing studies and how the writing process may contribute to L2 development. It is believed that one of the best ways for students to notice gaps between
their output and target-like output is to provide them with teacher feedback (Bitchener & Storch, 2016). Study 2’s findings support the noticing hypothesis and skill acquisition theory, suggesting that sustained conscious output practice together with teacher feedback facilitates L2 writing development.

1.3 Thesis Chapter Overview

This thesis consists of ten chapters. Chapter 1 (this chapter) presents the background and establishes the motivation for the current study.

Chapter 2 begins with a theoretical discussion on L2 writing development, providing a brief overview of some theories employed in L2 writing research. Two theories—Schmidt’s (1990) noticing hypothesis and DeKeyser’s (2007) ‘practice’ in skill acquisition theory, which informed the study and were used to explain the findings from the study—are also discussed. The chapter then reviews various cross-sectional and longitudinal studies on writing development. These reviews are followed by a discussion of the factors affecting writing development. The chapter then presents the studies related to background learners. It concludes with my identification of the current research gap and the presentation of this project’s three main research questions.

Chapter 3 briefly discusses aspects of the Japanese writing system. It then provides a review of the literature on measures for assessing writing development. This chapter highlights and discusses in detail the measures deemed most relevant to Study 1.

Chapter 4 establishes the methodology for Study 1. It first describes the context of the study and the instruments used. The chapter describes the pilot study undertaken prior to Study 1, and then presents a detailed description of the data-coding procedures used for the pilot study. It also describes the statistical procedures used to analyse the data. The chapter ends by presenting the inter-coder reliability analysis for Study 1.
Chapter 5 presents Study 1’s results, together with a discussion. The results and discussion answer the first RQ: that is, whether or not the writing ability of L2 Japanese learners at a university differs across five proficiency levels.

Chapter 6 establishes the methodology for Study 2. Study 2 implemented a classroom-based longitudinal design. The chapter begins with an overview of the context of the course undertaken by participants. It then provides information on the study design, followed by the data collection tools and data-coding procedures. The chapter concludes with an overview of the reliability analysis and data analysis procedure.

Chapter 7 presents the results based on text analyses for Study 2. Collectively, the results and analysis address the second and third research questions—how writing develops in the short (over one semester) and long term (over one year), and the difference in writing development between character-based learners and non-character-based learners. Both the short- and long-term findings are presented. Finally, the development of character-based learners and non-character-based learners will be compared.

Chapter 8 reports on the findings from the semi-structured interviews conducted with all participants. It first presents the results from the student interviews, followed by the results from the teacher interviews. The objective of the interviews was to determine participants’ perceptions of writing development, as well as whether or not classroom activities and assessment tasks assisted in improving students’ writing.

Chapter 9 discusses the findings of Study 2. It relates the findings to previous studies that have investigated L2 writing development. The findings are also discussed in light of the two SLA cognitive theories (the noticing hypothesis [Schmidt, 1990], and ‘practice’ in skill acquisition theory [DeKeyser, 2007]).
The thesis concludes with Chapter 10. This chapter first provides a summary of the entire project’s findings, followed by a discussion on the theoretical, methodological and pedagogical implications of these findings. Finally, Chapter 10 acknowledges some limitations of the current research project and suggests directions for further research.
Chapter 2: Literature Review

Japanese is one of the most widely studied foreign languages in Australia. Despite the fact that teaching writing at an early stage is considered important for character-based languages such as Japanese and Chinese (Cook & Bassetti, 2005), little research attention has been given to writing development in these languages. Most studies on second language (L2) writing development have been conducted in an ESL/EFL context. A small body of literature has investigated L2 Japanese writing (e.g., Aoki, 2010; Iwashita & Sekiguchi, 2009; Machida, 2001); however, most of these studies have focused on writing performance at one point in time and not on writing development. Thus, L2 Japanese writing development remains under-explored. Moreover, most studies (in an ESL/EFL context, as well as languages other than English) appear driven by pedagogical questions, with an absence of theoretical frameworks that inform the studies’ research design and explain their findings.

This chapter provides a brief overview of the various theories that may explain L2 writing development. It also introduces the two L2 acquisition (SLA) theories that have informed the current study. This is followed by a review of the literature pertaining to L2 writing development. Two common approaches to investigating L2 writing development are examined: cross-sectional and longitudinal. The chapter will also discuss studies that have investigated students’ perceptions of writing development, followed by the four main factors that can affect L2 writing development. These are instruction, feedback, motivation and the learner’s background. The chapter concludes by identifying the research gap in the literature and establishing the research questions.

2.1 L2 Writing Development: Theoretical Perspectives

A number of theories can be used to explain how and why L2 writing develops. Manchón (2012) discussed the different theories that can inform L2 writing development research. These include dynamic systems theory, sociocultural theories of
language learning, theories of multi-competence in language learning studies, goal theories in education and psychology and genre theories in L2 writing research, and systemic functional linguistics. However, it is important to note that L2 writing is a very complex phenomenon, and no single theory may ever explain it (Cumming, 2010). As noted by Manchón (2012), ‘[t]he development of L2 writing capacities is intrinsically a multifaceted phenomenon that is mediated by a wide range of varied personal and situational variables’ (p. 5). The following section briefly discusses theories that inform L2 writing development.

### 2.1.1 Dynamic systems theories

Dynamic systems theories (DST) include dynamic usage-based (DUB) theory and the dynamic model of multilingualism (DMM). These theories attempt to explain change in complex systems. Even though DST has been recognised in science disciplines such as mathematics, neurology and psychology for some time, it was not until the 1990s that this theory was applied to SLA (Larsen-Freeman, 1997). Interest in the application of DST to SLA has grown since then (De Angelis & Jessner, 2012; Larsen-Freeman, 2002). It has also been applied in research on writing development in L1 and in additional languages.

In DST, the term ‘dynamic’ implies that the current level of development is connected with the previous ones and hence that ‘initial conditions’ are critical (Verspoor & Smiskova, 2012). Moreover, some scholars (e.g., de Bot, Lowie & Verspoor, 2007; van Geert, 1991) suggest that every element and sub-system is interconnected, constantly affecting each other during development, with every new stage emerging from the previous stage.

DUB theory applies DST to language learning. The term ‘usage-based’ originates from Langacker’s (1987) assumption that the basis of one’s linguistic system stems from usage events. Ellis (2002) argued that although frequency itself is not a
sufficient explanation for L2 development, it is a vital component in language development, as learners acquire that which they hear (i.e., input) and use frequently. DUB theory also assumes that internal factors, such as a learner’s motivation, attention and cognitive ability, affect language development along with external factors (Verspoor & Smiskova, 2012).

DUB theory suggests that language is emergent because it develops in complex, adaptive and active ways (Ellis, 2008). Language is formed through personal relations, when learners interact with others through their prior experience and present perception of various language forms (Verspoor & Smiskova, 2012). Verspoor, Lowie and van Dijk (2008) suggested that, as language consists of several interacting subsystems, development (including writing development) could experience moments of instability at any time. This instability is referred to as ‘variability’. In L2 development, this variability is highest when the learner is trying new strategies and structures. Thus, DUB theory suggests that differences should be expected among L2 learners, and will result in diverse individual trajectories.

According to DUB theory, to understand L2 development stages, it is important to not only examine how each of the subsystems develops over time, but also investigate how they interact. In other words, it is important to understand the relationship between various components of the system. For example, to produce longer and more complex sentences, a language learner will need to understand a greater number of complex words (Verspoor et al., 2008).

Several studies have applied DUB theory to examine L2 development and the possible interactions between various variables (e.g., Caspi, 2010; Spoelman & Verspoor, 2010; Verspoor et al. 2008). For example, Verspoor et al. (2008) examined the development of English in advanced Dutch learners, particularly the relationship between the learners’ lexicon and syntax in academic texts. The study revealed
compromises between more varied word use and longer sentences at different developmental stages. Spoelman and Verspoor (2010) examined beginning Dutch learners as they acquired Finish and investigated development in relation to learners’ use of various complexity measures at the word, noun phrase and sentence level. These three variables developed simultaneously; as the learner’s word complexity increased, so did the noun phrase and sentence complexity.

DMM extends DST to multilingualism (De Angelis & Jessner, 2012; Herdina & Jessner, 2002; Jessner, 2008). In DMM, a multilingual system is dynamic and adaptive and is ‘characterised by continuous change and non-linear growth’ (Jessner, 2008, p. 273). Further, De Angelis and Jessner (2012) argued for a number of factors being involved in multilingual development—development not only originates from social, psycholinguistic and individual factors, but also from various learning environments (natural, instructional or various combinations of both).

Multilingual development is more complicated than just learning multiple languages. For example, De Angelis and Jessner (2012) examined the interaction between three languages known by participants (Italian [L1], German [L2] and English [L3]) and how such interactions and relative proficiency levels influenced writing development in additional L2s over time. They showed that a complex relationship existed between the three languages and proficiency in each of these languages.

2.1.2 Sociocultural theory

Sociocultural theory (SCT) is a major theoretical approach to understanding human cognitive development. Although SCT is not an L2 development/learning theory, it is increasingly recognised as a useful theoretical framework for L2 teaching and learning. Drawing on Vygtosky (1978), SCT is a psychological theory that explains the development of complex human cognitive skills.
The principle assumption of SCT suggests that human cognitive functions, such as attention, planning and problem solving, are mediated mental activities but that the sources of these cognitive functions are social activities (Swain, Brooks & Tocalli-Beller, 2002). According to Vygotsky, interactions between the novice and the expert or more competent other (e.g., a teacher or peer) can contribute to the novice’s development from other-regulation to self-regulation. In other words, the novice learner will be able to undertake tasks and solve problems independently through the assistance of a more knowledgeable person. Nevertheless, it is important to note that the effectiveness of this assistance will depend on its quality as provided by the expert. The expert must consider the novice’s current level of knowledge and the potential level that can be attained through assistance (Wells, 1999). Vygotsky (1978) used the term ‘zone of proximal development’ (ZPD) to distinguish these two levels of performance (current and potential). Thus, effective assistance is considered an important factor in ZPD. This (i.e., effective assistance) is referred to as scaffolding. Scaffolding can be defined as an instructional assistance by the more competent other to help novice learners achieve tasks during a learning activity (Wigglesworth & Storch, 2012). It is worth noting that Vygotsky did not restrict mediation during ZPD only to experts, but suggested that peers could also contribute to each other’s development if they worked cooperatively. Hence, learners with similar L2 proficiency can still provide effective assistance to each other, as seen in a number of studies on collaborative writing (e.g., Storch, 2013; Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009, 2012). For example, Wigglesworth and Storch (2012) demonstrated that not only did peer feedback during collaborative writing provide students with ‘opportunities to learn’ through pooling their linguistic knowledge, but also enhanced their writing development.
2.1.3 Multi-competence theory

Writing development can also be explained through multi-competence theory, which suggests that knowledge of an L1 and L2 language are not separate; rather, they are a partially merged system (Kobayashi & Rinnert, 2012). Multi-competence was developed by Cook (1991, 2002, 2013) and is defined as the ‘knowledge of more than one language in the same mind or the same community’ (Cook, 2013, p. 1). Cook (2008) argued that the mind of a multilingual learner is not the same as the mind of a monolingual learner or the sum of two minds of monolingual learners. The mind of a multi-competent learner has overlapping brain storage sites, with a shared lexicon in L1 and L2.

The notion of language transfer in multi-competence theory is useful when examining writing development. Of particular interest to this thesis is Kobayashi and Rinnert’s (2012) study. Drawing on multi-competence theory, Kobayashi and Rinnert (2012) investigated the transferability of text features across L1 (Japanese) and L2 (English) using data from a series of studies (see Table 1.1). Table 1.1 summarises the focus, types of research and references to relevant studies by Kobayashi and Rinnert.

Table 2.1

Studies on Transferability of Text Features across L1 and L2

<table>
<thead>
<tr>
<th>Stage</th>
<th>Main focus</th>
<th>Types of research</th>
<th>Studies</th>
</tr>
</thead>
</table>
| Stage 1  
(N = 27) | Transfer from L1 to L2 | Cross-sectional within-subject comparison | Rinnert & Kobayashi (2007) |
| | | | Kobayashi & Rinnert (2008) |
| | | | Kobayashi (2005) |
| | | Case study | |
| Stage 2  
(N = 36) | Transfer from L2 to L1 | Cross-sectional within-subject comparison | Kobayashi & Rinnert (2007a, 2007b) |
| | | | Rinnert & Kobayashi (2009) |
In order to illuminate the ways in which multi-competence theory can inform the understanding of writing development, Rinnert and Kobayashi (2012) focused on three main features when analysing L1 and L2 essays produced by their study participants: the types of discourse, subtypes of argumentation, and introductions and conclusions. They argued that these features represented cross-rhetorical differences and evidence of transfer across languages. The results from cross-sectional, cross-linguistic and longitudinal comparisons suggest that novice writers depend on the L1 knowledge they have acquired. In contrast, the most experienced EFL writers demonstrate a sophisticated knowledge of introduction components and genre-related rhetorical features, as well as greater audience awareness in both L1 and L2. In addition, Rinnert and Kobayashi (2012) suggested that when writers became more experienced, some L1/L2 knowledge was combined into a single component with greater degrees of overlap. Thus, the authors concluded that multi-competence theory not only assists in understanding how L2 writers develop their L1 and L2 writing knowledge over time, but also how their L1 and L2 writing merges and becomes interrelated.

2.1.4 Goal theory

Goal theory is also considered crucial when understanding L2 writing development, as students’ goals for writing have a considerable impact on their development. In turn, this influences the way in which educators should intervene most effectively to improve their students’ abilities (Cumming, 2012). Goal theory

<table>
<thead>
<tr>
<th>Stage</th>
<th>Main focus</th>
<th>Types of research</th>
<th>Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 3</td>
<td>Identification of L1 text features</td>
<td>Cross-linguistic comparison</td>
<td>Rinnert, Kobayashi &amp; Katayama (2010)</td>
</tr>
<tr>
<td>(N = 40)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 4</td>
<td>Development of L1 and L2 writing ability</td>
<td>Longitudinal study</td>
<td>Kobayashi &amp; Rinnert (2010)</td>
</tr>
<tr>
<td>(N = 1)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Kobayashi and Rinnert (2012, p. 107)
(Cumming, 2012) views writing development from a different perspective: ‘[d]evelopment is operationalized as control over the production of L2 texts, and the emphasis is on the antecedents of behaviour that contribute to that control’ (Norris & Manchón, 2012, p.225). Goal theory suggests that students’ writing goals reflect not only their development of literacy, but also their orientations to knowledge and the world. This creates the potential for effective instructor intervention to improve students’ abilities.

Two types of goals exist, according to goal theory: performance and mastery goals. Performance goals are simply oriented to perform and complete a task. Mastery goals also involve task performance, but extend to gaining better control over learning (i.e., attaining a certain skill). Cumming (2012) suggested that, depending on the type of goal, learning trajectories could differ. Moreover, Cumming (2012) argued that these two goals were not only the focus for an individual’s long-term writing development, but that they also offered the potential for teaching and learning opportunities to influence learners’ development.

Cumming (2012) used goal theory to compare different groups of learners in Toronto to determine the effect of different goal orientations on L2 writing development (for a more detailed description of the study, see Section 2.5.4 on motivation and L2 writing). Cumming’s (2012) study found that goals for L2 writing differed depending on the learner’s socioeconomic status, and that this could affect a learner’s writing development.

2.1.5 Genre theory

Genre theory is significantly and directly relevant to writing development. According to Hyland (2007), genre refers to ‘abstract, socially recognised ways of using language’ (p. 149). Common academic and professional writing genres include research articles, letters of recommendation, research presentations and theses (Tardy, 2012).
The assumption behind genre theory (and by implication a genre approach to instruction) is that by developing an awareness of genres, learners will be able to improve their writing skills as they will have a better understanding of the ways in which language structures are used in different contexts, for different purposes. In genre pedagogy, purpose and audience are considered the two most important variables that enable writers to perform social actions (Pasquarelli, 2006).

Empirical studies on genre learning in relation to writing development are relatively scarce (Tardy, 2012). However, those scholars (e.g., Hyon, 2001; Tardy, 2005, 2012; Yasuda, 2011) who have examined writing development through genre teaching suggest that teaching approaches focusing on rhetorical organisation can improve L2 learners’ writing. Tardy’s (2012) two-year longitudinal study in the United States of America (US) with four advanced L2 ESL writers from different backgrounds (i.e., a different country of origin and discipline) demonstrated that genre knowledge development is important for understanding writing development. The researcher investigated learners’ genre knowledge through analyses of their own writing, feedback from mentors, oral reflections on their writing and the texts and tasks with which learners were engaged. The study found that learners tend to use common strategies for L2 genre learning and that different writing contexts will enable different learning resources (Tardy, 2012). In addition, Tardy (2012) argued that it was important for researchers to consider developments in genre knowledge when investigating L2 writing development, especially in academic or professional contexts.

Genre theory can also be examined from a systemic functional linguistics (SFL) theory perspective. SFL primarily draws on Michael Halliday’s work in the 1960s and sees grammar as a social action of meaning driven by the functions and purposes of a community (Ortega, 2009). SFL considers the relationship between meaning and form, context and text, content and wording, and views these pairing as inseparable. It
explains how people interpret experience through meaning making (Ortega, 2009). Halliday’s SFL theory in social semiotics has influenced many traditions of genre theory. It is especially linked to genre theory in Australia. Genre is defined in SFL as the ‘forms of language and the social settings that shape the language’ (Zisselsburger, 2011, p. 52).

From this perspective, the structure of language serves three sociocultural purposes. SFL identifies these as ideational, interpersonal and textual. The ideational functions to transmit information and indicates what is happening in a text (Schleppegrell, 2004). The interpersonal establishes social relations and connects a society’s members; it can be thought of as taking a stance, with the relationship between the writer and reader being communicated through this stance (Schleppegrell, 2004). In order for writers to communicate effectively to their intended audience, they must make language choices, consider context and determine how to convey their stance. Textual choices provide relevant discourses suitable for the context (Schleppegrell, 2004).

Several scholars (e.g., Achugar & Columbi, 2008; Byrnes, 2012; Yasuda, 2011) have investigated writing development by applying SFL to their study. For example, Yasuda (2011) examined the development of genre awareness, linguistic knowledge and writing competence of Japanese undergraduate students (N = 70) learning English. Yasuda referred to SFL when in defining genre. Email writing was chosen as a writing task based on SFL theory. The results indicated that students improved in their genre awareness over time (as revealed by interview data), which was also reflected in their written emails. Students’ writing improved significantly in terms of task fulfilment and appropriateness, cohesion and organisation, grammatical control, fluency and language sophistication.
The above theories have informed L2 writing development from different perspectives. The current study is informed by two cognitive SLA theories. These are introduced in the following section.

2.2 Theories Informing the Current Study—L2 Writing Development from Second Language Acquisition Cognitive Perspectives

SLA theories have been used in the past mainly to explain L2 learning processes and the development of oral skills. L2 writing scholars (e.g., Bitchener & Storch, 2016; Williams 2012) have suggested in recent years that certain SLA theories and theories of cognitive development can also inform L2 writing studies. Within the framework of cognitive psychology, two SLA theories inform the current study. These are the noticing hypothesis and the idea of ‘practice’ in skills acquisition theory.

2.2.1 Noticing hypothesis

Schmidt (1990) proposed the noticing hypothesis as a cognitive concept in SLA. The basis of this concept stems from Schmidt’s (1983) findings in an earlier study, where he found that: 1) the frequency of a language form in the input does not result in acquisition unless it is noticed by the learner, and 2) the corrective feedback of a learner’s errors does not contribute to learning unless noticed by the learner. The noticing hypothesis suggests that any language form should be noticed in the input and registered consciously for acquisition to result (Schmidt, 1990, 2001). Three different levels of consciousness are defined in the noticing hypothesis (Schmidt, 1990):

- **Level 1: Perception**—forming internal representations of external events but with perceptions that are not necessarily conscious.

- **Level 2: Noticing (focal awareness)**—having an awareness of a certain aspect of the environment selected by the learner.

- **Level 3: Understanding**—attempting to comprehend the significance of the object that has been noticed and then analysing and comparing it with what
has been noticed on other occasions. This level of consciousness also involves thinking, problem solving and meta-cognition.

According to Schmidt (1990), ‘the subjective experience of noticing is the necessary and sufficient condition for the conversion of input to intake’ (p. 132). In other words, it is vital for learners to notice intake consciously. Schmidt suggests that this level of noticing is necessary for the learning process in acquiring various aspects of language. Noticing is defined as being aware of the gap between a learner’s output and what is required to be produced (Schmidt, 2001). It is possible to close this gap through intervention with a teacher’s feedback (Swain, 1993).

A number of studies (e.g., Bitchener, 2012; Bitchener & Knoch, 2010; Van Beuningen, 2010) support the notion that teacher feedback through form-focused intervention assists students with noticing gaps in their SLA process. This then leads to writing development in the sense of increased grammatical accuracy. In the input process, teacher feedback can increase the salience of target language structures, directing learners’ attention to their linguistic gaps. Without teacher intervention, learners may not notice these gaps. According to Van Beuningen (2010), it is easier for learners to notice their gaps through corrective feedback on their writing rather than on oral production, and this will promote SLA.

It is important to note that learning requires more than noticing. According to Schmidt (1990), learning goes beyond noticing; it depends on a learner’s insight and understanding. Unless learners have the opportunity to engage in tasks that use what they have been taught, it may be difficult for learning to occur. The following section discusses how practice can assist learners to engage in tasks that promote L2 development.
2.2.2 Notion of ‘practice’ in skills acquisition theory

Fitts and Posner (1967) were the first to introduce a three-phase model of skill acquisition. This included a cognitive, an associate and an autonomous phase. Developing this concept, Anderson (1982) later proposed an adaptive control of thought-rational (ACT-R) theory. This theory suggests that three stages are involved in the human production system: declarative, procedural and automatic. DeKeyser’s (1997) skills acquisition theory was based on ACT-R, positing that L2 acquisition occurs through these stages. DeKeyser (1997) suggested that learners encountered similar stages of development during the three stages of skill acquisition theory: an early stage of knowledge representation, behavioural change and finally, fluency in the second language.

The notion of practice in DeKeyser’s (2007) skills acquisition theory explains how prolonged systematic practice can lead to automatised knowledge. Learners develop through three stages when acquiring a skill through practice. In the first stage, learners obtain information and acquire declarative knowledge. For example, the acquisition of declarative knowledge may occur when a learner has been taught new grammar features through explicit teaching. In this stage, the learner must retain what has been taught in the working memory using a considerable amount of attention. As such, a significant portion of the working memory’s capacity is required (DeKeyser & Criado, 2013). The second (procedural) stage requires the learner to turn the acquired knowledge into behaviour; that is, the learner must practice what has been learned. If the learner uses their declarative knowledge appropriately, proceduralisation can be achieved through repeated practice. For example, a learner can practice their declarative knowledge through guided composition, gap filling and substitution exercises, turning it into procedural knowledge. In the last stage, for the newly acquired knowledge to be automatised, extensive practice is required to decrease the learner’s reaction time, error
rate and interference from other tasks (DeKeyser, 2007). Hence, fluency and accuracy may increase through meaning-based practice due to declines in reaction times and error rates. This quality of automatisation is known as the ‘power law of practice’. The more that procedural knowledge is automatised, the more fluent and accurate it becomes; nevertheless, the ‘power law of practice’ predicts that subsequent automatisation will be slower.

2.3 Approaches to Investigating L2 Writing Development

Two research designs are commonly used when investigating L2 writing development: cross-sectional and longitudinal. In a cross-sectional design, data from different groups with different proficiency levels are compared. Although such studies cannot observe individual development, they allow researchers to collect and compare a relatively large amount of data (Mann, 2003). In longitudinal studies, although it is possible to examine individual development, collecting data can be challenging, as this requires studying learners over an extended period. The following sections first discuss cross-sectional studies, and then longitudinal studies. Most of these studies do not state explicitly which theoretical framework informed their research; as such, those studies that are informed by a particular theory are discussed accordingly.

2.3.1 Cross-sectional design

Cross-sectional designs are useful in determining patterns for a certain group. They can identify associations that can then be studied thoroughly using other methods, such as longitudinal studies (Mann, 2003). The literature on cross-sectional studies examined below is divided into two sections: 1) studies in an ESL/EFL context and 2) studies of languages other than English.
2.3.1.1 Cross-sectional studies in an English as a second language and English as a foreign language context

To gain a better understanding of L2 writing developmental stages, researchers have increasingly focused on the relationship between writing performance and L2 proficiency (e.g., Cumming et al., 2006, Sasaki & Hirose, 1996). Isolating the different variables may explain the differences in L2 learners’ ability. Many scholars (e.g., Cumming et al., 2006, Sasaki & Hirose, 1996) have investigated the textual aspects (vocabulary, fluency and accuracy) that distinguish high-proficiency and low-proficiency learners by 1) grouping students according to test scores and 2) conducting discourse analyses of textual features. Examining discourse measures such as fluency (i.e., the number of words), accuracy (e.g., grammatical accuracy) and complexity (e.g., type token ratios measuring vocabulary variation within a text) allows researchers to investigate what particular aspects of writing develop over time (these measures are discussed in detail in Chapter 3). Cumming (1989) suggested that ‘as people gain proficiency in their second language, they become better able to perform in writing in their second language, producing more effective texts, and attending more fully to aspects of their writing’ (p. 121). This is supported by Sasaki and Hirose (1996), who studied the expository writing of 70 Japanese university students (majoring in British and American studies) with low- to high-intermediate English proficiency. Proficiency levels were based on English composition scores from the Comprehensive English Language Test for Learners of English (CELT). The writing composition was rated on five criteria: content, organisation, vocabulary, language use and mechanics. L2 proficiency had a significant influence on students’ L2 writing product: writing fluency determined the differences among students in terms of writing performance. In other words, the higher the proficiency level, the more words a learner wrote.
Similar to Sasaki and Hirose’s study (1996), other scholars (e.g., Banerjee, Franceschina & Smith, 2007; Cumming et al., 2006) have also used large-scale proficiency tests (such as the Test of English as a Foreign Language [TOEFL] and the International English Language Testing System [IELTS]), to investigate L2 writing performance and compare essays produced at different proficiency levels. For example, Cumming et al. (2006) analysed the discourse features of six prototype written tasks for the new TOEFL at three different proficiency levels (Levels 3, 4 and 5), investigating 216 written compositions by 36 test-takers. Texts were analysed in terms of lexical and syntactic complexity, as well as grammatical accuracy, to examine language use. The researchers not only examined the use of language, but also analysed texts in terms of argument structure, orientation to source evidence, and the verbatim use of sources. Cumming et al.’s (2006) results showed that the more proficient the learner, the better she or he performed in all measures. Test-takers with higher-proficiency levels tended to use a wider range of words and more clauses; they also demonstrated greater grammatical accuracy and summarised evidence sources more effectively (Cumming et al., 2006). This result was consistent with Sasaki and Hirose’s (1996) study, which also indicated that L2 proficiency had a significant influence on students’ writing performance.

In a more recent study, Knoch, Macqueen et al. (2014) investigated the discourse features produced by test-takers at different scoring levels for the independent and integrated writing task in the TOEFL internet-based test (TOEFL iBT) writing section. This study replicated and extended Cumming et al.’s (2006) study, which used the TOEFL independent and integrated prototype tasks (the integrated task in their study differed from the TOEFL iBT). The new integrated task consisted of a reading-to-write task and a listening-to-write task, which required examinees to write a summary based on reading and listening to a text. Knoch, Macqueen et al. (2014) analysed the
discourse features of 480 test-takers’ written compositions, using the measures of accuracy, fluency, complexity, coherence, cohesion, content, orientation to source evidence and metadiscourse. Significant differences appeared in discourses produced at different score levels and measures. For example, test-takers with greater writing ability produced essays that were more accurate. However, the results for syntactic complexity were mixed; the number of clauses per T-unit increased with higher-proficiency learners for the integrated task, but did not increase for the independent essay. Results for the cohesion measure were unexpected; learners at higher-proficiency levels used fewer anaphora words than did learners at the lower proficiency levels.

Similarly, Banerjee et al. (2007) investigated the writing performance of 275 test-takers in Tasks 1 and 2 of the academic writing module in IELTS Bands 3 to 8 (Bands 1, 2, 9 being unavailable). Task 1 required test-takers to write a summary of at least 150 words in response to a particular graph, table, chart or process. With Task 2, test-takers were given a certain topic and required to write at least 250 words discussing the topic. Banerjee et al. (2007) included L1 background (writers with Spanish or Chinese backgrounds) and task-type (Tasks 1 and 2) as independent variables; the discourse analytic measures used included accuracy, complexity, cohesive devices and vocabulary in the writing samples. Banerjee et al. (2007) found that all measures increased with improved proficiency levels, except for syntactic complexity. The L1 and task-type variables affected some measures, and the tasks affected vocabulary richness in various ways. For example, Task 1 scripts tended to be more lexically dense than Task 2 scripts.

The cross-sectional studies discussed above suggest that in most cases, L2 proficiency has a significant influence on learners’ writing performance—significant differences were seen in the discourse produced at different proficiency levels and measures. These studies were conducted in either an ESL or EFL context. The
following section reviews cross-sectional studies in the context of a language other than English.

2.3.1.2 Cross-sectional studies of languages other than English

Many previous cross-sectional studies have focused on learners of English; however, a few studies have investigated learners’ proficiency levels and writing performance in languages other than English. Examining writing performance in languages other than English is important, as language structures differ. Accordingly, the way in which learners perform in one aspect (e.g., accuracy and complexity) may differ from one language to another. This may also influence a learner’s writing development. Three different languages—Spanish, Russian and Chinese—are discussed in the first section, followed by Japanese, which is directly related to this study.

Henry (1996) argued that many foreign language teachers in the US follow the L2 writing development standards guidelines listed in the American Council on the Teaching of Foreign Languages (ACTFL) Proficiency Guidelines. Researchers such as Valdes, Haro and Echevarriarza (1992) and Henry (1996) examined the validity and reliability of these guidelines with learners of Spanish (Valdes et al., 1992) and Russian (Henry, 1996). For example, Valdes et al. (1992) examined the writing development of Spanish learners (N = 38) from a private US institution at three different levels:

1. those in the second quarter of their first year (Level A)
2. those in the fourth quarter of their second year (Level B)
3. those in an advanced composition class (Level C).

The study had two aims: 1) to determine any differences in writing abilities, depending on the proficiency level and 2) to confirm whether the L2 Spanish students would follow the writing development sequence as outlined in the ACTFL Proficiency Guidelines. In-class writings on the general topic ‘Me’ (students were asked to write about themselves) produced by students from each level were collected and analysed.
using three different criteria: 1) quality of message 2) organisation and style and 3) standard of language of use. A clear distinction in these criteria was evident between the different levels. The study also revealed that learners did not follow the ACTFL developmental sequence when they began to write in Spanish. The first stage of the ACTFL developmental sequence indicates that foreign language (FL) learners usually begin by writing only familiar words and phrases. However, the researchers suggested that learners did not start from zero, but applied the writing abilities they had acquired through English when writing in Spanish. Valdes et al. (1992) concluded that a clear relationship existed between L1 and L2 writing ability, especially when both languages shared the same alphabetic writing system. Hence, when examining learners’ L2 writing development, this study suggested it was important to consider if the learners’ L1 and L2 were similar. Moreover, it was also vital to examine learners from backgrounds with dissimilar languages, as the developmental sequence might differ yet again.

Henry’s (1996) study on Russian learners at the University of Iowa (US) examined writing development in a language that was more dissimilar to English than was Spanish. Sixty-seven autobiographical essays were collected from four different levels of Russian language study: 20 from students in their first semester, 19 from students in their second semester, 19 from students in their fourth semester and nine from students in their sixth semester. Essays were analysed statistically in terms of fluency (length of essay), syntactic fluency (T-unit length) and accuracy (percentage of correct T-units and number of correct T-units per 100 words). Only fluency and syntactic fluency differed in terms of performance. Regarding essay length (fluency), students in their second, fourth and sixth semesters showed greater fluency compared to first-semester students. Results for syntactic fluency showed a clear distinction among the groups, with first-semester students scoring the lowest and sixth-semester students scoring the highest. However, regarding accuracy, not only did the measures of
accuracy not distinguish between the different levels, but all types of errors (e.g.,
spelling, syntax, letter formation) were evident in students’ writings at each level. In
relation to the ACTFL guidelines, Henry (1996) concluded that these guidelines did not
provide a mechanism that differentiated between writing scripts for L2 Russian
learners’ writing development.

Examining the writing development of learners of languages even more distinct
from English than Spanish and Russian is vital. An example of such a language would
be Chinese, as it uses a logographic (pictorial) writing system in contrast to the
alphabet-based English, Spanish and Russian writing systems. Very few studies on L2
Chinese writing exist. However, Jiang (2012) recently conducted a cross-sectional study
on L2 Chinese learners (all native speakers of English and enrolled at the University of
Queensland, Australia) at three different proficiency levels: Level 1 (first year, n = 30),
Level 2 (second year, n = 53) and Level 3 (third year, n = 33). The study also included a
native control group (students from China n = 66). Students were given a letter-writing
task and 182 writing samples were collected. To examine L2 Chinese writing
development, this study used three T-unit measures: T-unit length, error-free T-unit
length and percentage of error-free T-units (EFTs). Hunt (1970) proposed the T-unit and
defined it as:

The shortest units in which a piece of discourse can be cut without leaving any
sentence fragments as residue. Thus, a T-unit always contains just one
independent clause plus however many subordinates clauses are attached to the
independent clause. (p. 188)

Jiang’s (2012) study found that the percentage of EFTs was the only measure
that discriminated between different proficiency levels. This study’s outcomes differed
to those of Henry’s (1996), where accuracy (the percentage of correct T-units and the
number of correct T-units per 100 words) revealed no differences in the scores between
different levels in L2 Russian learners. The measures that distinguish between different proficiency levels may differ depending on many factors, such as context, tasks, students’ language background or the way the levels were established. This is evident in the three studies mentioned above; for example, the levels established differed slightly from study to study. The above literature suggests that language (e.g., Chinese and Russian) can also be a factor. T-units and other measures will be discussed in more detail in Chapter 3.

Similar to L2 Chinese, few cross-sectional studies have been undertaken on Japanese. In one of the few studies undertaken, Ishige (2007) conducted a cross-sectional study on L2 Japanese learners with a native Korean background studying at a language school in Japan. Ishige chose Korean participants, as they comprise the second-largest population studying Japanese in Japan, after the Chinese. Moreover, in recent decades, Koreans have stopped incorporating Chinese characters into their writing system. This means that current generations of Korean students are not as familiar with kanji as previous generations. Ishige wanted to investigate the writing development of these non-character-based learners. Participants were from three different levels: upper-beginner \((n = 27)\), intermediate \((n = 26)\) and upper-intermediate \((n = 21)\). At the time of Ishige’s study, upper-beginner students had completed approximately 280 hours of study, intermediate students had completed 380 hours of study and upper-intermediate students had completed 520 hours of study. The details of instruction at the language school were not available. Participants were given 40 minutes to write an essay and the data were analysed in terms of content, organisation and use of language (i.e., vocabulary, use of conjunctions and use of clauses). Ishige found that development was evident in all areas (i.e., content, organisation and use of language). However, differences in the scores for organisation and use of language were greater than differences in the scores for content. Ishige argued this was because to
write and express one’s self clearly (i.e., with content), the organisation and use of language must improve before development relating to content can occur. This is especially so for those at the upper-beginner and intermediate levels. Therefore, Ishige posited it was important to teach students about the use language and organisation so they could express themselves more effectively before they improved their writing’s content quality.

The studies examined in the above sections show that cross-sectional designs are useful to identify learners’ developmental changes. Further, large-scale proficiency tests, such as TOEFL and IELTS, enable comparisons between different proficiency levels using large amounts of data. This is evident in Knoch, Macqueen et al. (2014) and Banerjee et al.’s (2007) studies. Nevertheless, the major drawback with cross-sectional studies is that they do not enable any conclusions to be made about individual development. Thelen and Smith (1996), among others, have questioned whether all subjects in a cross-sectional design follow the same developmental pathway. Cross-sectional designs may not necessarily obtain accurate data, as development varies from individual to individual. This is particularly the case if individuals come from different backgrounds. Developing advanced L2 capacities takes time, especially in relation to writing, as learners must have both L2 proficiency and general writing skills. The following section discusses studies that have investigated writing development using longitudinal studies.

2.3.2 Longitudinal design

Longitudinal studies on L2 writing development have usually been conducted in an ESL/EFL context— with either pre-university students (e.g., Elder & O’Loughlin, 2003) or university students (e.g., Storch & Tapper, 2009; Xudong, Cheng, Varaprasad & Leng, 2010)—in relation to a specific course of instruction. Relatively few studies have been conducted on learners’ writing development without reference to a specific
course of instruction (e.g., Storch, 2009). Further, many of these studies were conducted over relatively short periods (e.g., over one semester). Longitudinal studies can encounter difficulties, as collecting data can be very expensive; it can be challenging to study learners over a long period (Myles, 2008). Nevertheless, a number of scholars have undertaken longitudinal studies over more than one year, mainly with university students in an ESL/EFL context (e.g., O’Loughlin & Arkoudis, 2009; Sasaki, 2007, 2009, 2011). Most of these studies used a pre-test/post-test design and employed various discourse analytical measures (i.e., fluency, accuracy and complexity) or band scores to examine writing development.

Most studies that investigated the impact of instruction on learners’ writing development have revealed that instruction has a significant influence on improving learners’ writing skills (e.g., Elder & O’Loughlin, 2003; Storch & Tapper, 2009). Studies that used English proficiency tests (such as IELTS) to examine writing development (e.g., Green, 2005; O’Loughlin & Arkoudis, 2009) found that initial scores were a stronger predictor of outcomes than the course length. However, scores based on average gains should be interpreted with caution, as learners with lower initial writing scores tend to improve more than did learners with higher initial writing scores. Instruction, especially corrective feedback, seems to be a critical factor in learners’ writing development, as accuracy can improve in as little as one semester (e.g., Storch & Tapper, 2009).

### 2.3.2.1 Studies on L2 writing development with/without instruction

A number of studies have examined the effect of instruction on writing. Some scholars have investigated whether or not writing develops without any formal English support (for studies in an ESL context). In this section, studies that included instruction will be discussed first, followed by studies without formal writing instruction (i.e., learners studying for degree courses at a L2-medium university).
2.3.2.1.1 Studies with instruction

A number of studies (e.g., Bae & Lee, 2012; Elder & O’Loughlin, 2003; Storch & Tapper, 2009; Xudong et al., 2010) have investigated the effect of instruction on L2 writing development, using tests such as IELTS. However, details about the instruction varied from study to study, and some studies provided insufficient detail.

Elder and O’Loughlin (2003) conducted one of the earliest studies on L2 writing development. The authors compared the IELTS scores of 112 pre-university ESL learners from four different institutions in Australia and New Zealand before and after an intensive 10- to 12-week English-language course. Four different proficiency levels were included: intermediate, upper-intermediate, advanced and upper-advanced. This study did not provide much detail regarding the nature of instruction. The authors noted that three out of four institutions used English for academic purposes (EAP) in their curriculum. Learners progressed during their course, with an average gain of approximately half a band overall. In terms of writing score gains, Chi-square analyses revealed that learners’ writing scores at the outset (the lower the better) and their perception regarding the importance of writing for future academic study emerged as the best predictor of writing gain. For those learners with lower English-language proficiency (ELP), improvements were greater than were those with higher ELP. Although significant improvement was seen among learners, Elder and O’Loughlin indicated that the scores in their study were based on average, not individual gains, and so suggested the results should be interpreted with caution, as individual performances may vary.

Other scholars have conducted similar studies, but with university-level students. For example, Storch and Tapper (2009) examined the writing development of 69 international postgraduates (Masters and PhD) concurrently enrolled in a semester-long (12-week) EAP course, Presenting Academic Discourse (PAD), alongside their
degree program (PAD is designed specifically for postgraduate students). The instruction focused on developing skills such as critical reading and writing, research proposal writing and appropriate academic language, with an emphasis on structure, accuracy and academic vocabulary. The participants were divided into two groups: 18 high-proficiency learners (IELTS score of 6.5 or above) and 34 low-proficiency learners (IELTS score of 6) (17 participants whose English scores were not reported were omitted). A comparison of the learners’ pre- and post-tests found that text structure and rhetorical quality improved for most students. Storch and Tapper suggested that as students had frequent exposure to academic texts in their concurrent courses, this may have influenced their writing. As opposed to many other studies that found no improvement in linguistic accuracy (e.g., Shaw & Liu, 1998; Storch, 2009; Xudong et al., 2010), Storch and Tapper (2009) concluded that grammatical accuracy improved for both low- and high-proficiency learners. Significantly, in relation to Storch and Tapper’s study, PAD teachers provided feedback on writing (mainly on structure and language). The amount of feedback provided in the other studies (Shaw & Liu, 1998; Storch, 2009; Xudong et al., 2010) that included instruction was unknown. Also unknown was whether learners had to revise their writing following feedback, as was the case in PAD. Storch and Tapper’s (2009) results thus highlight the importance of feedback for improving students’ writing, especially in relation to accuracy.

Longitudinal studies on writing development are usually conducted with university-age students. Studies involving younger learners are relatively scant, with one notable exception. Bae and Lee’s (2012) study involved 42 EFL students in Korea over 18 months (approximately 315 hours of instruction), with participant ages ranging from 9 to 13. The students were enrolled in an EFL program called English for Young Learners (EYL), and were taught by a native English speaker after school. Class size was smaller than mainstream daytime classes. In terms of writing, students practiced in
specific genres, such as narrative, descriptive and persuasive writing. Although major assignments were due at the end of semester, students had the opportunity to write two or three drafts, on which they received corrective feedback. Students were also given homework for every class. Regular corrective feedback was given for vocabulary, spelling, grammar and discourse. To assess students’ writing development, three tests were conducted at nine-month intervals across the 18 months (a picture-based story-writing task was used for all three tests, but each test had a different picture). The five components that measured their English-writing abilities were as follows: grammar, content, coherence, spelling and text length. The results showed significant improvement by all students. Grammar improved at a similar rate to coherence, but the rate of improvement in relation to text length was much higher than all other components. To assess the development of the EYL students’ achievements further, Bae and Lee (2012) compared the five components with native fourth-graders from the US. Although EYL students had a lower mean score for the grammar component than US fourth grade students, they had a higher mean score for spelling and text length. This indicates that even if learners do not live in the target language (TL) country, with proper instruction and a favourable learning environment, it is possible to improve their writing ability in a reasonable period and even achieve better outcomes in some aspects than the TL country’s native speakers. Drawing on the notion of ‘practice’ in skill acquisition theory (DeKeyser, 2001), Bae and Lee (2012) suggested that providing students with writing practice through various activities leads to improvement. In addition, as with many other L2 writing scholars (e.g., Ferris, 2003; Storch & Tapper, 2009) Bae and Lee (2012) also stress the significance of feedback for L2 writing development.

The studies discussed above were all undertaken in ESL/EFL contexts; as mentioned, few studies on Japanese exist. However, some scholars have attempted to
investigate the writing development of L2 Japanese learners through longitudinal studies (e.g., Iwashita & Sekiguchi, 2009; Tanaka, 1997). Tanaka (1997) investigated how Japanese writing skills can develop through in-country experience in Japan. The study involved ten Australian secondary school students, ranging from Year 10 to 12. Nine out of the ten students had studied Japanese (for a minimum of 2 years and a maximum of 6 years, depending on the student) prior to visiting Japan as exchange students (the learners’ initial proficiency was not mentioned). All ten students spent one year at a Japanese secondary school and lived in Japanese homes. The study did not provide curriculum details. Compositions written by the ten participants (pre- and post-Japanese visit) were examined for writing development. Various linguistic measures were used to analyse the data: fluency, accuracy, complexity, general conjunctions, clause conjunctions and several sentence-ending variables. Students improved in all measures between their pre- and post-Japanese visit compositions, except for words per error-free sentences (categorised as ‘accuracy’ in Tanaka’s study), which showed a decline. Students made the most improvement in the number of words used (fluency). It is interesting to note that although students’ writing improved after they had spent one year in Japan, some used a greater number of colloquial expressions in their post-Japanese visit compositions. Tanaka argued this was due to the influence of spoken language—students had many opportunities to converse with Japanese people during their stay. Hence, in-country experience had both positive and negative influences on the students’ writing.

The above studies examined learning undertaken with a structured form of instruction. The results suggest that factors such as explicit instruction, feedback and frequent practice improve students’ writing. Some writing scholars have investigated whether learners’ writing develops without specific instruction. The following section reviews these studies.
2.3.2.1.2 Studies without instruction

Several studies have investigated L2 writing development undertaken without specific instruction. For example, Storch (2009) investigated L2 writing development in an immersion context. The study examined the L2 writing development of ESL university students, but these participants did not receive any formal English-language support during the semester. Twenty-five international students participated in the study. The participants’ English proficiency was equivalent to an IELTS score of 6.5, which is the minimum threshold required for entry into this particular university. A test-retest design was used to investigate the learners’ developmental changes. After one semester, students improved mainly in the structuring and development of ideas. Some improvement was also seen in language formality; however, no improvement was evident in accuracy or complexity. Given that the learners did not attend EAP courses during the semester, Storch concluded that simply attending normal classes and having frequent exposure to academic texts did influence learners’ writing development. However, this may not be as effective as a structured form of instruction during semester. This was also evident in studies that involved specific instruction (e.g., Bae & Lee, 2012; Storch & Tapper, 2009). Storch and Tapper (2009) suggested that one reason for students not improving their accuracy and complexity might be attributable to the short duration of study (one semester) and the absence of feedback.

Other studies have examined L2 writing development over the course of one year or more (e.g., Knoch, Rouhshad et al., 2014; Knoch, Rouhshad, Oon & Storch, 2015), concluding that a longer time spent at an institution did not necessarily lead to greater improvements in writing. For example, Knoch, Rouhshad et al. (2014) examined the writing development of 101 ESL students over one year in an English-medium university in Australia. All participants had met the university’s English-language criteria upon entry (i.e., an IELTS overall score of 6.5 or equivalent). Approximately
half of these students had received their secondary education in Australia and only a small proportion of participants (11) took EAP during the first year of their study. The Diagnostic English Language Assessment (DELA) writing task was used to assess the learners’ writing development. The same essay topic was given after one year, and essays were scored for three criteria: organisation, content and form. Knoch, Rouhshad et al. (2014) not only compared DELA writing scores, but also examined different discourse analytic measures: accuracy, fluency, grammatical complexity and lexical complexity. Students’ writing only improved in terms of fluency.

In a more recent study, Knoch et al. (2015) investigated learners’ L2 writing development over three years. Thirty-one undergraduate L2 students at a large university in Australia participated. Similar to Knoch, Rouhshad et al.’s (2014) study, participants had met the university’s English-language entry requirement (i.e., an IELTS overall score of 6.5 or equivalent). A test-retest design was used: participants were given 30 minutes to write an argumentative essay on the same topic at the beginning and the end of their degree. The study found that students only improved in terms of fluency over the three-year period. No significant changes were evident in terms of accuracy, grammatical and lexical complexity. Semi-structured interviews were conducted to determine students’ L2 writing opportunities, as well as how much feedback they had received. From the student interviews, it was evident that writing requirements over the three years of their study were minimal. Further, the interview results revealed that when students did submit a writing assignment, they received no feedback. Knoch et al. (2015) indicated several possible reasons for this lack of improvement: the relatively short length of the writing task, writing topics not being discipline-specific, few writing opportunities and minimal feedback. As noted in the section on studies with instruction, the authors here also emphasised the importance of ‘practice’ and ‘feedback’ in L2 writing development.
Writing, when compared to other areas of language learning (i.e., listening, speaking and reading) may be the last area in which students improve when given limited or no instruction and minimal writing practice. This was evident in O’Loughlin and Arkoudis’s (2009) study. These authors examined the change in IELTS scores over the duration of undergraduate and postgraduate studies of 63 international students (30 undergraduates and 33 postgraduates from the Faculty of Economics and Commerce) in an Australian university. Study participants were not enrolled in EAP courses.

O’Loughlin and Arkoudis (2009) compared the learners’ original IELTS results (used for entry into the program) with another test taken midway through their final semester. Undergraduates were about to finish their three-year program, while postgraduates were about to finish their 12- or 18-month program. This meant that the time between the first and second tests differed for undergraduates and postgraduates. Writing showed the least improvement of all the language skills examined. O’Loughlin and Arkoudis (2009) suggested this might be because, depending on the subjects selected, the students were not required to complete extended pieces of writing. As such, they did not have enough writing practice during their course of study. Regression analysis indicated that the learners’ initial writing scores were the best predictor of improvement in writing scores. Previous studies had achieved the same results (e.g., Green, 2005; Elder & O’Loughlin, 2003); this aligns with the hypothesis that greater improvement is seen in learners with lower initial scores compared to those with higher initial scores.

As the above studies (e.g., Bae & Lee, 2012; Storch & Tapper, 2009) indicate, instruction helps learners to improve their writing when compared to no instruction. Thus, this literature review indicates it is important to undertake learning with specific instruction in relation to L2 writing development. However, learners’ perceptions about whether their writing improved or not may differ to those shown through quantifying
and comparing pre and post-test results. These perspectives may also help explain why L2 writing develops (Storch & Tapper, 2009; Xudong et al. 2010). As such, these perspectives are important to consider. The following section reviews studies involving student perspectives of L2 writing development.

2.4 Students’ Perceptions of Writing Development

Students’ perceptions provide an important perspective in studies investigating writing development. Qualitative research seeks to uncover participants’ perspectives and understandings of the world. With an understanding of students’ perspectives, educators and researchers can determine whether a writing program is useful for L2 writing development (Xudong et al., 2010). Although few in number, such qualitative studies (e.g., Storch & Tapper, 2009; Xudong et al., 2010) reveal that students’ perceptions of writing development correspond to the focus of the L2 writing course they may have completed. In addition, students’ perceptions are linked to whether or not their perceived improvement corroborates the findings from the quantitative analysis of their writing performance (pre- and post-test).

For example, Storch and Tapper’s (2009) study (discussed earlier) investigated students’ perceptions of whether their EAP course had assisted their academic writing development. Participants ($N = 69$) either had an IELTS score of 6 or 6.5 at the outset of the study. The aim of this particular EAP course was to develop learners’ structuring, accuracy and vocabulary through explicit teaching in seminars. The course required students to compose a range of writing tasks, and provide feedback on the writing tasks and course materials. This was reflected in students’ questionnaire comments, where they reported the course had helped to improve their academic writing, particularly in structuring ideas and texts, grammar and vocabulary. This also corroborated with the researchers’ quantitative results, which revealed that learners mainly improved in accuracy, vocabulary and structure. However, it was not clear from the questionnaire
findings which specific aspect of the course (i.e., feedback, seminars and course materials) had assisted learners’ writing development.

Xudong et al.’s (2010) study, consisting of participants with a similar proficiency level as in Storch and Tapper’s (2009) study (i.e., IELTS 6 or 6.5 equivalent), also found that students thought the EAP course had contributed to their improved academic writing. Xudong et al. (2010) also reported on students’ perceptions of difficulties in writing. The three most frequently mentioned difficulties reported by students were choosing appropriate words and expressions, organising their ideas, and grammar. Moreover, although most students felt that the EAP course had helped with their writing, many reported that the course did not help to improve their grammar skills. These difficulties were reflected in their pre- and post-test comparisons, where no improvement in grammar was evident. Thus, the study suggests that students may be well aware of which aspects of writing are improved, and which are not.

Factors apart from instruction, such as sustained and challenging writing activities, can also influence students’ writing development. This was evident in Manchón and Roca de Larios’s (2011) longitudinal study, where students perceived that improvement in their writing originated from both instruction and the writing activities required over the year. Eighteen advanced EFL learners enrolled in an EAP writing course at a Spanish university participated in Manchón and Roca de Larios’s (2011) study. The data consisted of in-depth semi-structured interviews collected at two points in time, self-reflection journals, writing data (a standardised proficiency test and an argumentative essay at Time 1 and Time 2). The quantitative results obtained from the writing data revealed that participants improved in accuracy, fluency and lexical variety. Students reported they had improved mainly in terms of grammar and vocabulary. Their perceptions did not align exactly with the quantitative results; however, these did align with the classroom observation data, where vocabulary concerns dominated lessons.
Hence, regular exposure to new vocabulary in class helped students improve in that area. In addition, instruction and writing activities promoted an awareness of certain learning processes. For example, in the interviews, students reported they became aware of the gap between what they knew and what they were able to do. Many students felt that these learning processes (i.e., noticing the gap and determining how to deliver the intended message appropriately) allowed them to improve their writing.

Other studies (e.g., Knoch, Rouhshad et al., 2014; Storch & Hill, 2008) have shown that feedback was particularly important for writing development. For example, in Knoch, Rouhshad et al.’s (2014) study (discussed earlier), most students reported that they received no feedback on assignments from their lecturer. If (or when) they did receive feedback, it was not specific, but generic, such as ‘you need to improve your writing’, which did not assist their writing development. Apart from feedback issues, students also reported various grammatical and lexical difficulties when they had to write under pressure.

A number of participants in Storch and Hill’s (2008) study also reported they did not receive enough feedback during the semester and that opportunities to write were scarce. Storch and Hill correlated the results of their student questionnaire with the results from their discourse data. Interestingly, accuracy was influenced by learners’ integration with the host culture; the fact that students had native-speaker friends helped them improve their accuracy in writing. Further, those students who reported no difficulties with their English scored higher on all band scores than did other students.

The studies discussed above suggest that students perceive instruction (e.g., an EAP course), practice and feedback as important in their writing development. However, students’ perceptions of writing development can vary from individual to individual, due to a host of individual and context-related factors, such as their primary discipline and social aspects. This was apparent in Morton, Storch and Thompson’s (2015) study.
Morton et al. (2015) investigated the perceptions of three multilingual students concerning their academic writing in their first year at a large Australian university. The authors used a case study approach. Participants had all met standard English-language requirements for university entrance (an overall 6.5 out of 9 on IELTS, or equivalent). Data included background questionnaires, five semi-structured interviews and three writing tasks. All three participants improved their writing over the year; however, their perceptions of writing development differed. A student’s discipline was the main factor that contributed to differences in beliefs about writing development. For example, one participant (male) enrolled in an engineering degree was not required to undertake much writing and therefore felt he made no improvement in his writing. The authors also identified social factors as contributing to participants’ writing development. For example, one participant (female), who was not confident with her writing in the beginning, mentioned that becoming active and seeking outside help gradually (e.g., friends, flatmates and online contacts) assisted her with improving her writing and also increased her confidence. The researchers concluded that considerable variation existed in the students’ perceptions of what practising ‘academic writing’ meant. Some saw it as an issue of skills development, while others saw it as an issue of interpersonal relations or negotiating identities. This suggests that when investigating learners’ perceptions of L2 writing, many factors can contribute to writing development.

The studies discussed above show that students’ perceptions of writing development tend to correspond to the quantitative findings regarding accuracy, vocabulary and structure. Questionnaires and interviews indicate that instruction, practice in writing and feedback are the main factors identified as leading to or inhibiting L2 writing development. It is worth noting however, that studies on students’ perceptions of writing development remain scarce. Hence, the findings’ generalisability is not clear.
The studies reviewed earlier, as well as these studies on students’ perceptions, indicate that many factors can affect writing development. The following section discusses some of these.

2.5 Factors that Affect Writing Development

A number of factors can affect writing development. This section focuses on four main factors identified in the literature: instruction, feedback, motivation and learner background.

2.5.1 Instruction

That instruction plays a pivotal role in students’ writing improvement is well supported empirically (e.g., Bae & Lee, 2012; Storch & Tapper, 2009; Xudong et al., 2010). The issue then is what kind of instruction is most appropriate. A number of approaches can be incorporated in L2 writing classes to assist students improve their writing. Four major approaches will be discussed in this section: product, process, genre and process-genre approaches. Each will be explained by focusing on what teachers do in class (i.e., classroom and student activities), as well outlining the underlying philosophy of each approach regarding how L2 writing is learned and developed.

2.5.1.1 Product approach

The product approach emphasises the learner’s finished product. Writing is considered a form of linguistic knowledge, focusing mainly on the use of vocabulary, syntax and cohesive devices (Pincas, 1982). Explicit teaching of grammar is considered important and the main criteria for good writing in this approach are accuracy and clear exposition (Hyland, 2003b). Pedagogically, teachers often use rigidly controlled instruction methods, such as guided composition, gap filling and substitution exercises (Hyland, 2001). These tasks consist of learners imitating, copying or transforming models provided by the teacher. The underlying assumption of this approach to L2 writing instruction is that writing development results from imitating and manipulating
models given by the teacher (Hyland, 2003b). Learners are likely to improve their grammatical accuracy through this approach. Further, Shannon (1994) argued that by using a grammatically correct and structurally organised model, students would learn to write well in any rhetorical mode.

Theoretically, the product approach reflects the notion of ‘practice’ in skill acquisition theory. Drawing on this theory, the acquisition of declarative knowledge occurs when the learner has been taught a new grammar feature through explicit teaching. L2 learners have the opportunity to practice their declarative knowledge through guided composition, gap filling and substitution exercises, turning it into procedural knowledge. Having learners undertake practice (with activities such as drilling) may assist with the gradual automatisation of grammatical knowledge.

Although the product approach allows learners to develop their writing through imitating and practicing sentence patterns and structures, a number of disadvantages are identified. One criticism levelled at the product approach is that it is teacher-centred: a method in which students do not have the opportunity to discuss, negotiate or interact with each other or the teacher (Mourssi, 2013; Zhou 2015). This approach only offers opportunities for students to imitate fixed patterns mechanically. Further, teacher feedback often only focuses on the use of language (e.g., grammatical errors). However, Cross (1991) claimed that these features work well for beginners and are often better suited to learners with lower proficiency. This is because these features facilitate vocabulary development, sentence structure knowledge and self-confidence.

Many linguists (e.g., Silva, 1990; Zamel, 1983) argue that the product approach focuses too heavily on the product of writing. Instead, they promote a process approach to writing instruction. The process approach is discussed in the following section, which highlights the underlying differences between the two approaches. This next section
also discusses what is perceived to lead to writing development under the process approach.

2.5.1.2 Process approach

As opposed to the teacher-centred product approach, the process approach is student-centred—it examines what writers do in the composing process to find new meanings (Lam, 2015). With the product approach, students do not have the opportunity for peer-to-peer discussion. In contrast, with the process approach, students can discuss and negotiate freely with their peers and teachers and receive feedback. Emphasising the composing phases (i.e., planning, drafting and revising), as well as opportunities to communicate with their peers and teachers, is assumed to help students improve their writing skills. Shannon (1994) argued that, as this approach allows students to choose their own topic (in contrast, topics are chosen by teachers in the product approach), students would be more motivated to write. Thus, motivation may play an important role in L2 writing development—this factor will be discussed subsequently.

The main role of teachers in this approach is to guide students during the writing process by assisting them to develop strategies for generating, drafting and revising their ideas (Hyland, 2003b). Teachers use pre-writing activities, encourage brainstorming and outlining, request multiple drafts, give extensive feedback, facilitate peer response and delay surface corrections until the end (Raimes, 1992). Teacher and peer responses are regarded as crucial factors during the process for students to develop their writing skills. Hence, teachers provide feedback in various forms: teacher-student conferences, audiotaped feedback, peer response and reformulation (Hyland, 2011).

The underlying assumption in this approach is that submitting multiple drafts and receiving feedback throughout the writing process will lead to writing development, as students are given the opportunity to correct and improve their writing during the
process. The process approach assumes that students will develop their writing skills, as teachers are constantly attending to the writer (Liebman-Kleine, 1986).

The process approach was adopted by ESL writing teachers in the 1970s and 1980s. However, a number of scholars (e.g., Horowitz, 1986; Hyland, 2003a; Reid, 1994) criticise this approach, identifying several pitfalls. First, this approach depends heavily on revision and does not consider the importance of training students in timed tasks, such as essay writing under examination conditions (Horowitz, 1986). Second, this approach assumes universality; it does not consider that each individual is different. Process writing may be effective for some but not for others (Reid, 1994; Shannon, 1994). Third, students do not have the opportunity to consider how different texts are written for particular purposes and audiences (Hyland, 2003a).

Hyland (2003b) noted there is little evidence that the process approach on its own leads to any significant improvement in writing. He argued that many factors affect the writing process. Learners need assistance regarding how to write and in understanding how texts are constructed by topic, purpose, audience and cultural norms. This criticism has led to interest in the genre approach in recent years; this approach is described next.

2.5.1.3 Genre approach

The genre approach views writing as a social practice that connects writers and the world through genres (Johns, 2011). When something is written, this is usually done so with a purpose. To achieve this purpose, social conventions structure the message so that the purpose can be communicated to the reader. These abstract, ‘socially recognised ways of using language for particular purpose’ are referred to as genres (Hyland, 2003b, p. 18).

The main aim of this approach is to teach learners how to use language patterns that will enable them to write a purposeful and coherent piece of writing. To achieve
this objective, teachers focus on the text and content of a particular genre, for example in reports, editorials and business letters. The genre approach has three phases: 1) models of a certain texts are provided to students, 2) text construction is undertaken jointly by the teacher and students, and 3) a text is constructed independently by each student (Cope & Kalantzis, 1993).

Under the genre approach, teachers clarify what must be learned, provide a framework for studying both language and context, ensure that objectives are derived from the students’ needs and create appropriate resources for students (Hyland, 2007). Teachers use feedback as the main tool for scaffolding, which allows students to build their literacy resources and then participate in their target communities (Hyland & Hyland, 2006). This approach emphasises the need to become aware of textual and lexico-grammatical features associated with specific genres (Hyland, 2007). This awareness will help students to communicate effectively for a particular audience and thus improve their writing. Advocates of the genre approach believe that optimal learning occurs when learners are provided with explicit knowledge of the language and structural features of the texts they are required to write (Lee & Wong, 2014).

However, some scholars argue that (to a certain extent) the genre approach is an extension of the product approach (e.g., Dovey, 2010). As with that approach, teachers provide a model text and students imitate that model to improve their writing. Others argue that the genre approach undervalues the processes required to produce a text and considers learners passive (Badger & White, 2000).

To accommodate students’ needs and teach effectively, teachers often use a mixture of the three approaches discussed above (Hyland, 2003b). Badger and White (2000) suggested that the process-genre approach (a combination of the process and genre approaches) might be most effective for catering to students’ needs. This is discussed in the next section.
2.5.1.4 Process-genre approach

According to Badger and White (2000), incorporating the insights of the process and genre approaches is the most effective methodology for teaching writing. They named this the ‘process-genre’ approach.

The objective here is to teach students the relationship between purpose and form for a particular genre, while they use the recursive processes of pre-writing, drafting, revising and editing. Badger and White (2000) suggested that when learners lack knowledge in a language, three main sources of input could assist them: the teacher, other learners and examples from the target genre. Language-awareness activities using genre analysis (e.g., students learning the type of sentence structure and vocabulary used in a certain texts) are considered particularly important for learners’ development. Input on the skills required for writing (e.g., grammar construction, sentence structure) may also differ. In some cases, direct instruction from the teacher may be effective, whereas in other situations, demonstration by the teacher or other skilled writers may lead to a better result (Badger & White, 2000).

Four main approaches to writing instruction have been identified, but it is important to examine whether teachers are incorporating these approaches and if so, whether these approaches have been successful in improving learners’ writing. Previous studies examining writing development with instruction (discussed in detail in Section 2.3.2.1.1) did not indicate clearly what kind of approaches they incorporated. However, it appears that the process-genre approach is the most common. For example, the process-genre approach was used in both Storch and Tapper’s (2009) and Bae and Lee’s (2012) studies; these authors suggest that this approach is effective in improving learners’ writing. As discussed previously, the way this combined approach incorporates different methods ensures it can cater to learners’ different needs. One key element in these studies is the feedback learners received on their writing, whether from
teachers or peers. The following section discusses the studies on feedback related to writing development.

2.5.2 Feedback

Feedback has long been regarded as important for enhancing and consolidating learning in education; this can also be applied to the L2 writing context (Hyland & Hyland, 2006). According to Ferris (2003), feedback is critical for language learners to develop their writing skills. In the process approach, feedback is used as a developmental tool to enhance learners’ self-expression. In the genre approach, feedback is the main tool for scaffolding provided by the teacher, enabling students to build their literacy resources so they can participate in their target communities (Hyland & Hyland, 2006). Hyland and Hyland (2006) claimed that, although feedback was a central concern in L2 writing programs globally, the existing research was not explicitly positive about the role of feedback in L2 writing development. Issues related to feedback will be discussed in the following section. This section will first review the studies on teacher feedback, along with some studies on peer feedback. Finally, comparative studies on peer and teacher feedback will be discussed.

2.5.2.1 Teacher feedback

The benefits of feedback on students’ writing development are contentious. Feedback on the content and organisation of students’ writing is beneficial (e.g., Fathman & Whalley, 1990; Huntley, 1992; Kepner, 1991), providing that it is clear and unambiguous (Leki, 1990; Zamel, 1985). For example, Zamel (1985) examined the comments written by 15 teachers on 105 students’ compositions and found that many teachers provided vague responses, and often misread students’ texts. Further, they mostly corrected local errors and did not attempt to comment on the composition’s global meaning. These comments resulted in revised drafts that lacked coherence. Zamel (1985) suggested that teachers’ approach to feedback might be attributable to
teachers viewing writing as a ‘product’ rather than a ‘process’. Teachers saw themselves as language rather than writing teachers (in that they focus on the use of vocabulary and syntax).

Feedback on language use (corrective feedback) is debated vigorously in the field of L2 writing. Interest in the effectiveness of corrective feedback followed the debate between Truscott (1996, 1999, 2004, 2007) and Ferris (1999, 2004, 2006, 2010). Specifically, Truscott (1996, 1999, 2004, 2007) argued that error correction would not assist in writing development, as acquiring the forms and structures of a L2 was a gradual and complex procedure that error correction neglected. He also argued that empirical research on classroom studies that incorporated written teacher feedback (e.g., Kepner, 1991; Semke, 1984; Sheppard, 1992) revealed that written corrective feedback was counterproductive for the development of fluency and syntactic complexity. Moreover, error correction was ‘harmful’, as less time and energy was devoted to other productive aspects of writing. Theoretically, Truscott argued that corrective feedback played no role in the gradual inter-language development of L2 writers. He argued that corrective feedback might benefit explicit declarative knowledge at most, but not implicit and procedural knowledge, which is employed in the automatic use of language (DeKeyser, 2007).

Ferris (2004) presented an argument in support of corrective feedback. She noted that for adult L2 learners to continue their language development and avoid fossilisation, it was necessary to advise learners about their errors explicitly, and make their errors ‘salient’. A growing number of more recent empirical studies (e.g., Bitchener, 2008; Bitchener & Knoch, 2008, 2009, 2010; Fazio, 2001) have found that error correction is beneficial in improving students’ writing, especially in terms of accuracy. For example, Bitchener’s (2008) large-scale study ($N = 75$) in New Zealand (conducted with low-intermediate ESL students) found that targeted feedback, focusing
on two uses of English articles (referential indefinite ‘a’ and referential definite ‘the’) led to improved accuracy. The study found that students who had received corrective feedback outperformed those in the control group who had received no feedback.

Similar results were reported by Bitchener and Knoch (2010), in a study replicating Bitchener’s (2008) study, but with advanced ESL learners. Bitchener and Knoch (2010) concluded that focusing feedback on particular grammatical forms enabled advanced learners to improve their accuracy.

Ferris (2003) also noted that the efficacy of corrective feedback depended on how that feedback was provided. Teachers can provide corrective feedback in various ways: direct and indirect, focused and unfocused, coded and un-coded, or a combination of these. Direct corrective feedback occurs when the teacher indicates an error and gives the correct linguistic form. Indirect corrective feedback occurs when the teacher ‘indicates that an error has been made but leaves it to the student writer to solve the problem and correct the error’ (Ferris, 2002, p. 19). Indirect corrective feedback can be distinguished further through using codes. A coded error indicates the exact spot of an error, with an appropriate code (e.g., ‘SP’ for a spelling mistake and ‘T’ for a tense error). With un-coded feedback, the teacher only underlines or circles an error, leaving the student to identify and correct it (Bitchener, Young & Cameron, 2005). Ferris (2003) suggested that in terms of long-term writing development, indirect error feedback was more beneficial than direct error feedback, but that this also depended on student needs and the error type. In addition, Ferris (2003) suggested that focused error feedback (on a limited type of errors) was more productive than unfocused feedback (i.e., marking all errors). Providing extensive unfocused error feedback was not only time consuming for teachers, but also overwhelming for students. Finally, if teachers used coded feedback, it would be better to use a consistent system of codes so that students could readily interpret it (Ferris, 2003).
Research on the efficacy of these different corrective feedback methods has shown that learners who receive written corrective feedback (compared to those who do not) can produce accurate revisions and more accurate new written work immediately after receiving feedback. Further, they can often retain the knowledge gained from the feedback for a substantial length of time (e.g., Bitchener & Ferris, 2012; Bitchener & Knoch, 2010; Shintani & Ellis, 2013; Stefanou, 2014). Nevertheless, these scholars (e.g., Bitchener & Ferris, 2012; Shintani, Ellis, & Suzuki, 2014) admit that the number of error categories tested to date is limited. Hence, at this stage a firm conclusion cannot be drawn in relation to generalising the findings to all error categories (Bitchener & Storch, 2016).

A number of recent studies (e.g., Hubais & Dumaning, 2014; Rastgou, 2016), as well as the meta-analysis conducted by Biber, Nekrasova and Horn (2011), suggest that feedback combining a focus on form (corrective feedback) and content may be the most beneficial in terms of writing development. Biber et al.’s (2011) meta-analysis found that combining corrective feedback (focus on form) and content resulted in greater improvement compared to feedback that focused on form or content alone. The meta-analysis also revealed greater gains in grammatical accuracy were achieved when students received both types of feedback; however, no improvement in grammatical accuracy was evident when students received feedback on form alone. Hubais and Dumaning’s (2014) experimental study, where 17 EFL Omani students received form and content feedback, showed that learners improved their writing, both in terms of form and content. However, the study also revealed that students relied more heavily on form than on content feedback. Improvement in form was greater than improvement in content. Hubais and Dumanig (2014) concluded that although both types of feedback led to improvement, feedback related to form might facilitate better revision in
composition writing due to the tendency for learners to rely more heavily on this type of feedback.

One factor that may affect the usefulness of feedback is learners’ attitudes to the feedback they receive (Hyland & Hyland, 2006; Storch, 2010). The following section briefly reviews the studies about student perspectives of teacher feedback on L2 writing.

2.5.2.2 Students’ beliefs about teacher feedback

Students’ perceptions of the efficacy of feedback and their beliefs about what constitutes good writing may also influence the effectiveness of teacher feedback. It is also important that students and teachers have similar views about the efficacy of particular approaches to feedback and the type of feedback given (Diab, 2006).

Past research (e.g., Diab, 2005, 2006; Ferris, 1995, Leki, 1991; Zacharias, 2007) that investigated L2 students’ preferences and reactions to teacher feedback has suggested that students desire and expect corrective feedback from teachers. An early study conducted by Leki (1991) found that ESL students believed error-free writing constituted ‘good writing’. As such, teachers may lose credibility if not all surface errors are corrected. Zacharias’s (2007) study also revealed that students strongly preferred feedback on form, and considered this type of feedback as the most important factor in improving writing. This is because students thought that teachers were more competent in terms of language and knowledge.

Even though most surveys suggest that students prefer teachers to note their grammatical errors, some students also want feedback on content and ideas (e.g., Chen, Nassaji & Liu, 2016; Hedgcock & Letkowitz, 1994, 1996). Chen et al. (2016) investigated EFL learners’ perceptions of written corrective feedback. Overall, students showed favourable attitudes towards error corrections. However, the study also revealed that students strongly preferred extended comments on both the content and grammar of their written work.
Diab (2005, 2006) investigated teachers and students’ perceptions of various types of teacher feedback, reporting on some mismatches between teachers and students’ views on feedback. A university-level ESL instructor and two 19-year-old undergraduate international students participated in Diab’s 2005 case study. According to Diab, this particular teacher believed writing instructors should provide feedback on content rather than form but saw error correction as a useful ‘security blanket’ for students. The two ESL students stressed the relevance of comments and feedback in general and the importance of error correction and grammar in particular. Both the (one) teacher and (two) students in this study agreed that surface-level correction was necessary. In Diab’s (2006) much larger study, 14 EFL female teachers and 156 students from the American University of Beirut completed questionnaires in relation to preferences for error correction. Diab (2006) also revealed that a mismatch existed between the instructors and students’ preferences and their beliefs. For example, students preferred to receive error correction on their first draft, while teachers thought it more appropriate to note error corrections in the revised draft. As some of the teacher’s beliefs conflicted with the student’s views, Diab concluded that teachers should identify their students’ perceptions of feedback and minimise the gap between their own and students’ expectations.

The above studies on teacher feedback and students’ preferences for teacher feedback indicate that such feedback is important for L2 writing development. Alternative methods for providing feedback, such as through peers, can also help improve student writing. This is discussed in the next section.

2.5.2.3 Peer feedback

Peer feedback in writing involves students reading and commenting on other student’s writing. To assist students’ writing development, many teachers have incorporated peer feedback in classroom writing sessions; they believe that learners can
benefit from giving as well as receiving feedback (Lundstrom & Baker, 2009). Hansen and Liu (2005) argued that when applied appropriately, peer feedback can ‘generate a rich source of information for content and rhetorical issues, enhance intercultural communication and give students a sense of cohesion’ (p. 31). Past studies (e.g., Paulus, 1999) have shown that peer feedback can help students improve the vocabulary, content and structure of their writing. This means that peer feedback will not only help learners improve the local aspects of writing compositions, such as grammatical features, but also the global aspects, especially the overall structure and content (Ho, 2015). Lundstrom and Baker (2009) concluded that learners gained benefits from giving as well as receiving feedback.

The benefits of peer feedback are often investigated by comparing its impact in comparison to the impact of teacher feedback. The results of studies that have conducted such comparisons (e.g., Connor & Asenavage, 1994; Ruegg, 2015; Tsui & Ng, 2000; Yang, Badger & Zhen, 2006; Zhao, 2010) are mixed. Early studies in this area (e.g., Connor & Asenavage, 1994; Paulus, 1999) indicated that teacher feedback led to more revisions than did peer feedback (Connor & Asenavage, 1994; Paulus, 1999). Recent research (e.g., Tsui & Ng, 2002) indicates that peer feedback has other benefits that cannot be provided by teachers. For example, in Tsui and Ng’s (2002) longitudinal study conducted on secondary EFL learners in Hong Kong, both peer and teacher feedback assisted students’ writing development. Although students incorporated more teacher than peer feedback into their writing revisions, peer review had other benefits, such as allowing learners to gain a sense of audience. Further, it helped students become aware of the strengths and weaknesses in their own writing.

Other studies have found that peer feedback and teacher feedback may play different roles in students’ writing (e.g., Ruegg, 2015; Yang et al., 2006). Yang et al.’s (2006) study conducted in an EFL classroom in China shows some evidence of this.
Here, the researchers compared two groups: one received peer feedback only and the other received teacher feedback only. The researchers found that students who received teacher feedback produced mostly surface-level changes in their revisions, whereas more changes were apparent at the meaning-level from students who received peer feedback. Ruegg’s (2015) one-year longitudinal study also compared the impact on writing of a group that received peer feedback compared to a group that received teacher feedback. This study not only examined the types of changes in students’ revisions, but also the feedback given by peers and the teacher. The researcher found that significantly more of the teacher’s feedback was related to meaning-level issues and content, yet no significant difference was evident in the organisation, vocabulary, content or overall essay scores between the two groups, but the teacher feedback group performed better in grammar. Ruegg (2015) concluded that teacher feedback might be more effective for improving written grammar and that it might be better for teachers to provide feedback on grammar and content and for peers to provide feedback on organisation and content.

A number of factors, however, can influence the benefits of peer feedback. These include L2 proficiency, peer interaction, peer feedback training, and the learners’ perceptions and attitudes towards feedback—these are discussed below.

### 2.5.2.4 Factors affecting the benefits of peer feedback

A learner’s proficiency level must be considered when examining the impact of peer feedback. This is evident in Kamimura’s (2006) study. Kamimura examined the effect of peer feedback on learners with different English proficiency levels. Participants consisted of 12 high-proficiency and 12 low-proficiency Japanese university first-year students majoring in English. Students produced four essays (pre-test, post-test, original draft, rewrite). The findings revealed that both the low- and high-proficiency students improved their writing through peer feedback. Nevertheless,
greater improvement was seen in the high-proficiency students compared to the low-proficiency students. The study also found that highly proficient students tended to provide global comments in their peer feedback, while students with lower proficiency gave surface-level comments. This suggests that students’ proficiency levels affect the way in which they provide feedback; in turn, this affects the learners’ writing development.

Some researchers (e.g., Nelson & Murphy, 1993) have argued that the way in which students interact with each other during the peer review process can influence the uptake of peer feedback. For example, Nelson and Murphy (1993) conducted a study on four intermediate EFL learners from different backgrounds over six weeks. The researchers investigated the nature of interactions between students during the peer review process. The data consisted of video transcripts and the rough and final drafts of students’ written compositions. The findings showed that when students interacted with each other in a cooperative manner during the peer review process, they were more likely to incorporate their peers’ comments in their revision. Conversely, when students interacted with their peers in a defensive manner, or did not interact at all, the probability of incorporating their peers’ suggestions was low. Hence, Nelson and Murphy concluded that the way learners interacted with each other during the peer review process had a significant impact on the implementation of feedback. Peer interaction is important in the peer review process, but the outcome of the peer review can also depend on whether or not students have been given adequate training.

Empirical studies (Berg, 1999; Min 2005, 2006; Rahimi, 2013; Zhu, 1995, 2001) have indicated that training students through peer feedback-participation activities is vital in this process. These studies show that a trained peer response leads to a positive outcome in students’ revision type and quality. For example, Min’s (2005) study, which investigated the impact of trained responders’ feedback on EFL university students’
revisions, reported that training led to a significantly higher number of comments and a higher number of revisions (90% of all revisions). Further, the number of quality-enhanced revisions was higher than the revisions made prior to peer review training. Rahimi’s (2013) study revealed that through peer training, students were able to shift the focus of their feedback from only addressing formal errors, to comments that also addressed global issues (i.e., content and organisation). This helped students produce better quality writing.

Several studies (e.g., Fujieda, 2007; Stanley, 1992), have also found that training students in how to provide adequate feedback can provide other benefits, such as changing students’ perceptions about feedback. Fujieda (2007), who examined Japanese EFL writers’ \( N = 14 \) perceptions of peer feedback after they undertook prolonged training, found that most students thought peer feedback was a valuable tool for improving their writing. Students commented that peer feedback helped them in terms of grammatical errors and in relation to reconsidering vague arguments when revising their writing. Moreover, out of the 14 students, four who initially had a negative stance towards peer feedback developed affirmative feelings at the end of the course (the remaining eight had a positive attitude at the start, while two were neutral). By receiving instruction on how to provide feedback, students were able to communicate more effectively, engage more fully with the peer review task and outline clear and appropriate suggestions.

2.5.3 Reading and writing

Another factor that could affect learners’ writing development is reading. Krashen (1984) was an early proponent of the benefits of reading for writing development. Krashen (1984) argued that self-motivated reading for pleasure led to writing development. Grabe (1991) also emphasised reading’s importance as an input; through reading, learners increased their exposure to a language and hence acquired
new forms. Reading passages of text and using them as models for writing is an element of both the product and genre approaches to writing instruction. Several studies (e.g., Hafiz & Tudor, 1990; Lai 1993; Lee & Hsu, 2009) have investigated the reading–writing connection by examining the effect of extensive reading (ER) on writing in general, and on the acquisition of vocabulary in particular.

2.5.3.1 Reading and writing connections

Previous studies have suggested a correlation between reading and writing in one’s first language (e.g., Stotsky, 1983; Thorndike, 1973). For example, Stotsky’s (1983) survey on the L1 reading–writing relationship found that a correlation existed between reading achievement and writing ability—learners who were better at writing tended to be better at reading and vice versa. The amount of reading also reflected students’ writing quality. The study also revealed that correlations existed between measures of syntactic complexity in writing and reading ability. It appears that advanced readers can produce more syntactically complex writing compared to weaker readers.

However, research on the relationship between reading and writing in L2 has produced mixed results. Some of the early studies (e.g., Flahive & Bailey, 1993; Hedgcock & Atkinson, 1993) found no relationship between L2 reading and L2 writing. Hedgcock and Atkinson (1993) examined two groups: one consisting of 157 native English speakers and the other consisting of 115 ESL students. The study investigated L1 and L2 reading habits and writing through a self-reported questionnaire and scores on timed measures of learners’ expository writing proficiency. Although the results for L1 showed a significant relationship between reading and writing, no significant relationship was apparent between L2 writing scores and textbook or fiction reading in L2. Hedgcock and Atkinson (1993) suggested that either exposure to L2 texts had little
impact on L2 writing proficiency, or that the methodology was unable to detect any impact.

In contrast, some scholars (e.g., Janopoulous, 1986) have suggested that reading for pleasure in L2 and writing in L2 are closely related. For example, Janopoulous (1986) conducted a study on 79 ESL students to examine the relationship between L2 reading and writing. Participants produced a writing sample after one hour. In addition, they completed a questionnaire that revealed the amount of time spent for pleasure reading in both their native language and English. A 4-point rating scale was used to evaluate the writing samples holistically. The results showed a correlation between the proficiency levels for L2 reading and writing.

Other studies have shown that it is not only reading alone but ER that leads to better writing in both one’s first and second languages (e.g., Applebee, Langer & Mullis, 1986; Janopolous, 1986; Lee, 2001, 2005; Lee & Krashen, 1996). Moreover, ER is widely accepted as a method for language learning (e.g., Beglar & Hunt, 2014; Waring, 2006; Yamashita, 2013). Mermelstein (2015) defined ER as ‘reading as much as possible within the learner’s peak acquisition zone, for the purpose of gaining reading experience and general language skills’ (p. 182).

Considering that ER is advocated for language learning, many scholars (e.g., Elley, 1991; Hafiz & Tudor, 1990; Lai 1993; Lee & Hsu, 2009; Mermelstein, 2015; Tsang, 1996) have investigated whether adding ER into classrooms will improve EFL learners’ writing skills. Five of these studies in particular (e.g., Hafiz & Tudor, 1990; Lai, 1993; Lee & Hsu, 2009; Mermelstein, 2015; Tsang, 1996) promoted examining the effects of a reading program on the writing skills of EFL learners with similar educational backgrounds. These studies used descriptive writing to evaluate learners’ performance; the length of the reading programs ranged between 4 to 30 weeks.
The studies revealed positive results, with learners in reading programs showing more improvement than the comparison groups in several measures of language use. Out of these five studies, Lee and Hsu’s (2009) study reported the most improvement, with participants showing significant gains in six measures: fluency, content, organisation, language use, vocabulary and mechanics. However, the other studies did not show improvement in all measures; the treatment group in these studies outperformed the comparison group in several key areas (see Table 2.2).

Table 2.2

Comparison of Results on Studies Using Statistical Significance: Treatment Group

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Organisation</td>
<td>ns</td>
<td>sig</td>
<td>ns</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content</td>
<td>ns</td>
<td>sig</td>
<td>sig</td>
<td></td>
<td>sig</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>sig</td>
<td>sig</td>
<td>ns</td>
<td>ns</td>
<td></td>
</tr>
<tr>
<td>Language use</td>
<td>sig</td>
<td>sig</td>
<td>sig</td>
<td>sig</td>
<td></td>
</tr>
<tr>
<td>Spelling/mechanics</td>
<td>sig</td>
<td>sig</td>
<td>ns</td>
<td></td>
<td>sig</td>
</tr>
<tr>
<td>Fluency</td>
<td>sig</td>
<td>sig</td>
<td></td>
<td></td>
<td>sig</td>
</tr>
</tbody>
</table>

Note: sig = learners performing significantly better than comparison groups; ns = learners not performing significantly better than comparison groups

Even though the above studies demonstrate that positive results are gained through ER, it is also notable that some studies revealed no significant results (see Table 2.2). For example, Tsang (1996) reported that the treatment group did not perform better than the comparison groups in organisation, vocabulary and spelling or mechanics. Hafiz and Tudor’s (1990) study also showed negative results, with the treatment group having no significant differences on vocabulary. Lee and Hsu (2009) argued that studies prior to their own (i.e., Tsang, 1996; Hafiz & Tudor, 1990; Lai, 1993) had flawed designs that affected their results. The main weaknesses identified included the limited readings provided to participants, the short duration of the studies,
and that participants may have felt pressured by being required to write a report on what they had read (thus reducing their ‘pleasure reading’).

In an attempt to overcome these weaknesses, Lee and Hsu’s (2009) study was conducted over a longer period (30 weeks). It also included more reading material with less accountability. This resulted in a better outcome—as mentioned before, Lee and Hsu’s (2009) study showed the most improvement. Nevertheless, concerns about Lee and Hsu’s design exist. Mermelstein’s (2015) recent study examined the effect of ER on writing by again changing the design of previous studies. Mermelstein (2015) argued that Lee and Hsu’s study had problems with practicability, the treatment of ER and the requirement for participants to write a summary of their readings. The problem with practicability was that 50 minutes of uninterrupted reading time in a classroom was too long. Regarding the treatment of ER, although Lee and Hsu (2009) stated that learners in the experimental group could choose their own reading materials within their proficiency levels, it was unclear how the researchers determined the reading proficiency level of their participants. Hence, the participants may have chosen readings that were below their proficiency level. Finally, Mermelstein (2015) suggested that the requirement for participants to log their readings was too onerous.

To determine whether an even more enhanced ER design would produce a more positive outcome in learners’ writing skills, participants’ in Mermelstein’s (2015) study were placed according to their optimal reading levels. They experienced a longer and more balanced duration of ER, with ERs of between 15 to 20 minutes and had less accountability. As seen in the results (see Table 2.2), not all the measures showed significant differences between the treatment group and the comparison group in Mermelstein’s study.

In general, these studies provide evidence that ER does have a positive impact on writing development—ER can improve learners’ writing in terms of organisation and
structure, content, vocabulary, language use, spelling and fluency (Lee & Hsu, 2009; Shen, 2009). Shen’s (2009) study, conducted with first-year EFL college students and using students’ reading-log entries, creative writing and follow-up interviews, showed that reading also provided benefits in terms of developing critical thinking (and personal growth).

2.5.3.2 Vocabulary acquisition and extensive reading

Another well-documented benefit of ER is vocabulary acquisition. Vocabulary acquisition through ER depends on the number of encounters a learner has with each word and the quality of attention given to that word in each encounter (Nation, 2015). According to Nation (2001), three processes may lead to remembering a new word: noticing, retrieval and generative use. These processes are connected and the later stage is built upon the earlier stage. In each process, the quality will depend on whether the learner provides incidental or deliberate attention.

The first process in vocabulary learning is noticing; that is, learners must notice the new word (as with the noticing hypothesis). Incidental attention in noticing occurs when the learner’s focus is on aspects of communication other than that particular word, for example guessing from context. Conversely, deliberate attention occurs when the learner focuses on a particular word, such as finding a word in a dictionary.

Retrieval is the second process that leads to vocabulary learning. If the learner retrieves the word that has been ‘noticed’, then the memory of that word will be strengthened. When retrieving a word through incidental attention, this occurs through repeated reading of the same text. Deliberate attention for retrieval occurs when a learner puts a word encountered previously onto a word card and uses that card to learn the word (Nation, 2015). Repeated opportunities to retrieve the word are important. In other words, the time between a learner noticing a new word and retrieving that word is significant. It is difficult to strengthen the memory of the word if too much time has
passed between the previous and the present encounter with the word. This is because the present encounter will no longer be considered a repetition, but will be seen as a new encounter (Nation, 2001).

The third process in vocabulary learning is generative use. The generative process occurs when a learner uses or encounters a word that is used differently from the previous meeting with that word. For example, if the learner encounters the word ‘cement’ as a verb as in ‘we cemented our path’, but in the next encounter, the learner sees the same word in another context, such as ‘we cemented our relationship with a drink’. The learner will rethink the meaning of the word; this will assist in establishing the memory of that word (Nation 2001). Incidental attention in this process occurs when the learner encounters the word in a varied context while reading. Deliberate attention, on the other hand, occurs when the learner consults a dictionary while reading (Nation, 2015).

Through the three processes mentioned above, a learner is likely to remember a word. The more the learner encounters a word (i.e., repetition) the more likely it is that learning will occur. However, it is important to note that the quality of the meetings (incidental attention or deliberate attention) has a significant influence on the learning outcome. The deeper the quality of the meetings, the stronger effect they will have on vocabulary development (Nation, 2015).

2.5.4 Motivation and L2 writing

Motivation can play an important role in L2 writing development, particularly in terms of a student’s willingness to practice writing. Nevertheless, studies on motivation and L2 writing are relatively scarce, and much of the research on motivation has focused on SLA more generally (e.g., Dörnyei, 1994; Gardner, 1985; Yashima, Zenuk-Nishide & Shimizu, 2004). Two types of motivational orientation are described in SLA (these are also relevant to L2 writing): integrative and instrumental orientation.
Integrative orientation applies when an individual learns an L2 because of a desire to interact and engage with members of the TL community, while instrumental orientation applies when an individual is more concerned with L2 acquisition for pragmatic reasons, such as expanded employment opportunities. Of all the antecedents, ‘integrativeness’ has been most significant in the development of L2 motivation theory (Ortega, 2009). Kormos (2012) explained some of the potential roles motivation plays in L2 writing.

Kormos (2012) reviewed the role of individual differences in L2 writing, including the role of cognitive factors and the potential role of motivational factors. Kormos (2012) argued that writing is a time-consuming and complex task that requires determination and concentration. As such, a learner’s motivation and self-regulation are vital in determining whether learners will ‘engage in writing activities, what kind of writing task they will undertake, with what level of effort and attention they will approach the various phases of the writing process, and how they exploit the learning potential of [the] writing task’ (Kormos, 2012, p. 2). Due to the complex characteristics involved in a writing task, positive self-efficacy and interest become particularly important.

Interest and self-efficacy determine whether a learner engages in writing (Kormos, 2012). In other words, if students do not believe that they have the ability to acquire an L2, they will only complete the assigned writing task and are unlikely to engage further in cognitive processes (i.e., the problem solving behaviours that lead to L2 acquisition). Research suggests that students with a positive attitude towards writing tend to write more often and put effort into their writing tasks compared to those with a negative attitude (e.g., Hashemian & Heidari, 2013). For example, Hashemian and Heidari (2013) conducted a study on 30 Iranian EFL students, examining the relationship between the L2 learners’ motivation and their success in L2 writing.
Gardner’s (1985) attitude/motivation test battery questionnaire was used to determine whether students had a positive or negative attitude and whether they had an instrumental or integrative orientation towards the language. Hashemian and Heidari’s (2013) findings revealed a positive relationship between writing performance and an integrative orientation, but no such relationship was found in participants with an instrumental orientation. Further, students with a positive attitude towards SLA tended to write better quality texts than did those with a negative attitude.

Kormos (2012) argued that in order to carry out a certain writing task, both the motivation to write in L2, as well as the motivation to learn the L2 are needed. This is determined by various factors. The three main factors that influence a motivation to write are social, cultural and educational. These factors are important in forming the goals learners wish to achieve (Kormos, 2012).

For example, Manchón and Roca de Larios’s (2011) study showed how an educational factor (i.e., participating in an EAP course) influenced learners’ writing goals, which in turn led to their writing development. The study investigated the perceptions of language learning among 18 EFL learners at a Spanish university over nine months. Semi-structured interviews revealed that students’ goals for learning and writing changed over the course of the study—not only were students focusing more on accuracy, they had also started paying attention to text-type conventions and how they structured their ideas. While attending the EAP course, students received feedback and were engaged in demanding writing practice, which seemed to influence their writing goals. Thus, by expanding their writing goals, students were incorporating more ideas and writing longer, more linguistically complex texts that were also more coherent and cohesive. This was also seen in another study by Manchón (2009), where instruction had a positive influence on students’ beliefs, especially their sense of self-efficacy. This
suggests that instruction influences students’ motivation significantly, which then affects their writing development.

Similarly, a study conducted by Sasaki (2011) showed how social, educational and cultural factors affected students’ writing goals. Sasaki conducted a three-and-a-half-year longitudinal study on 37 Japanese EFL learners. Some had been studying abroad for different periods, while others had remained at home. The study revealed that students in the study abroad group improved their writing significantly compared to those who remained at home. One reason for this was that the writing goals of each group differed. Students in the study abroad group developed L2-related imagined communities, which motivated them to improve their writing, whereas the at-home group’s writing goals centred on obtaining a Bachelor of Arts degree. Sasaki’s study suggested that students who were intrinsically motivated tended to improve their writing more than those who were instrumentally motivated.

Sasaki’s (2011) study thus showed that different writing goals resulted in different writing outcomes. To determine the effect that different goal orientations had on L2 writing development, Cumming (2012) conducted a dual-phased study comparing two different groups of learners in Toronto: adult ESL students preparing for admission to university \((N = 45)\) and at-risk adolescents participating in an after-school tutoring program \((N = 21)\). Within goal theory, motivation is described as falling on two ends of a spectrum: performance-goal or mastery-goal orientation. Performance-oriented students only wish to fulfil minimum course requirements. Mastery-oriented students attempt tasks not just to achieve the highest mark possible, but also to gain knowledge and improve their abilities (Cumming, 2012). Through profile questionnaires and interviews, Cumming (2012) found that the ESL university students were highly motivated and had high-level goals. In contrast, the at-risk adolescents were not as motivated when asked to write in class, probably because they considered the purpose
of school activities was to ‘get good marks’, an objective they did not appear to enjoy or want to work towards. The conclusion drawn was that a learner’s goals, and hence their development in L2 writing, differed depending on their socioeconomic status.

The studies examined above (e.g., Cumming, 2012; Sasaki, 2011) suggest that motivation has an impact on students’ writing. Motivated learners tend to improve more than those who are less motivated (Cumming, 2012). Thus, motivation is one important factor when examining writing development.

2.5.5 Language background

In the context of learning a second language, the learner’s background (as well as goals and motivation) must also be considered. As noted in Chapter 1, students learning Japanese in Australia come from different language and cultural backgrounds. Speakers’ language backgrounds are distinguished in the literature in various ways (e.g., Elder, 1996, 1997; Scarino et al., 2011) and various terminologies are used. For the purpose of this study, terminology adapted from Elder (1997) will be used. These terms are defined below.

Background speaker (BS) learners are those with a background in a language similar to the TL they are learning. The Chinese and Japanese languages use many similar written characters, which may make the acquisition of Japanese easier for those from Chinese-speaking countries. Therefore, learners from China who are learning Japanese will be BS learners.

Non-background speaker (NBS) learners are those with no background in the language TL they are learning. Learners from Australia (native English speakers) without any relatives from a TL language will be NBS learners.

The learners’ backgrounds must be considered when testing and teaching NBS and BS learners in the same class. For example, in classes composed of both NBS and BS students, NBS students might feel disadvantaged. Studies that have investigated BS
learners and NBS learners of languages other than Japanese will be discussed here, followed by studies on BS learners and NBS learners of Japanese. Although the studies discussed in this section investigated the impact on performance rather than development, it is important to know how the performance of BS and NBS learners differs, as this can assist when examining writing development.

2.5.5.1 Background speaker and non-background speaker learners of languages other than Japanese

Elder (1997) produced an early study that investigated BS and NBS learners. Elder attempted to determine the difference in performance between BS and NBS learners of Italian, Greek and Chinese. This study was motivated by the large proportion of these nationalities in communities across Australia. Elder examined the assessment outcomes of students undertaking the Victorian Certificate of Education (VCE) in Languages Other than English (LOTE) for Year 12 BS and NBS learners in Australia in Italian ($n = 647$), Greek ($n = 667$) and Chinese ($n = 916$). Elder divided the study participants into four categories:

1. students with no exposure to the TL
2. students who used English at home but had some exposure to the TL
3. students who used the TL as the main language for communication at home
4. students who had been educated through the medium of the TL.

Although Elder investigated all four language skills (i.e., speaking, listening, reading and writing), only the results for writing will be discussed here. For both Italian and Greek, the outcome was surprising—Category 3 learners (i.e., BS learners) performed significantly worse than those in the other categories with a similar VCE writing task score. With Chinese, Category 4 students outperformed the other students by a significant margin. The result for Chinese students also revealed no significant difference between Categories 1 and 2, or between Categories 2 and 3. Elder noted that
differences in the results among the three languages might be due to the recent arrival (in Australia) of many BS Chinese learners, whereas BS learners of Italian and Greek tended to be second- or third-generation immigrants.

Although BS learners of Chinese outperformed NBS learners in Elder’s (1997) study, this was not the case in Ke’s (1998) study. Ke’s (1998) US-based study involved 85 BS learners and 60 NBS learners in their first year of university. Two tasks were used: character recognition (whether learners could recognise Chinese characters upon sight) and production recognition (whether learners could write Chinese characters sounded out to them). The results showed that language background did not have a significant influence on a learner’s ability to recognise and produce Chinese characters. One possible explanation for this result (i.e., no significant difference between the two groups) is that the non-background learners in this study may have been highly motivated learners who invested more time learning the Chinese characters compared to background learners. However, information on participants’ academic profile and their motivation for TL learning was unavailable; hence, Ke could not provide a reason to explain the differences.

Another study that examined the writing performance of BS learners studying Asian languages in Australia was the *Student Achievement in Asian Languages Education* (SAALE) project. This was a government-funded, two-year project that investigated the achievement of Years 6/7, 10 and 12 students in four Asian languages (Chinese, Indonesian, Japanese and Korean) (Scarino et al., 2011). The project aimed to determine students’ achievements in each language, focusing on the impact of a learner’s background on his or her achievements. Assessment tests for speaking, listening, writing and reading were conducted and marked by at least two judges. Questionnaires were used to identify each student’s background information and to divide students into different groups: 1) L1 learners (students who were born in, or
whose parents had come from, countries from the TL), 2) L2 BS learners and 3) L2 NBS learners. The main findings for writing performance related to the language background for all four languages were: 1) L1 learners consistently performed better than L2 learners, 2) the BS learners’ performance was variable; they were not as distinct from L1 and L2 learners, 3) BS learners performed better than did NBS learners. This suggests that regardless of the TL, L1 learners have an advantage over the other two groups and BS learners have an advantage over the NBS learners.

The above results (from SAALE) revealed that L1 learners consistently performed better. However, Kim (2012), who conducted a small-scale research project on the Korean language using the written component of SAALE data, argued it was important to examine the learners’ everyday language use and the habits of their L1, as results could differ considerably within the L1 learner group. Kim analysed the data of nine Year 10 participants from the SAALE project: five NBS learners and four BS learners. Of the four BS learner participants, two were born in Australia and the other two had arrived in Australia after two to three years of education in Korea: the ‘late arrivals’. Due to the small sample size, Kim was unable to make statistical comparisons between the two groups, but the findings showed variability in the outcomes within the BS group. BS learners born in Australia performed equally as well as the two late arrivals on the discourse task. Moreover, of the two learners born in Australia, the learner who had received less exposure to Korean at home performed better than the other learner who had been exposed to the language more often. Kim explained that the learner with more exposure to Korean at home seldom practiced writing in Korean. Thus, it can be said that having exposure to the TL language may not necessarily mean that a BS learner will perform well, especially in writing, if that learner does not practice frequently. This study reinforces the importance of ‘practice’ for L2 writing.
development. It is only through prolonged written output practice that learners can automatise their knowledge (DeKeyser, 2007).

Other than the SAALE project, which examined the achievement of Years 6, 7, 10 and 12 students, much of the previous comparative research has focused on learners at the tertiary level (e.g., Clyne et al., 1997; Elder 1997; Ke, 1998). However, Iwashita and Liem (2005) specifically examined BS and NBS learners among primary school students. Their study involved Year 6 students learning Chinese in Melbourne. Iwashita and Liem (2005) conducted tests in all four language skills (i.e., speaking, listening, reading and writing) to determine whether L1 would be less influential among BS and NBS learners at this level. They found that learners with an ethnic Chinese background achieved higher scores in all skills compared to learners with a non-Chinese ethnic background. It is worth noting that the differences between the two groups were particularly evident in writing and speaking skills. However, Iwashita and Liem’s (2005) study suggested that ethnic background was not the only factor that influenced BS learners’ high scores. For example, Chinese study outside school and the number of years of Chinese study at school affected test performance. Many BS learners receive extra tuition outside school to communicate with their relatives who do not speak English (Clyne et al., 1997).

2.5.5.2 Background speaker and non-background speaker learners of Japanese

Many studies examining the performance of BS and NBS learners of Japanese have focused mainly on reading or kanji comprehension. For example, Machida (2001) investigated the kanji vocabulary and reading comprehension skills of Chinese-background learners (n = 14) and non-Chinese-background learners (n = 14) studying Japanese at a university in Melbourne. Machida’s results showed that although Chinese-background learners scored higher in kanji comprehension, the non-Chinese-
background learners performed better at text comprehension. Machida suggested that having a good knowledge of *kanji* did not necessarily lead to a high level of text comprehension.

Similarly, Aoki (2010) conducted a study at the University of Adelaide to compare differences in *kanji* proficiency attainment between local students who were native English speakers and international students who had studied Chinese scripts in their home country. Sixty-six students enrolled in the second-year Japanese course participated in this study. Local students were divided into three categories: those who spoke English only, those who were bilingual in English and Chinese; and those who were bilingual in English and a language other than Chinese (e.g., Vietnamese). International students were also divided into three categories: Chinese, Korean and Malay. Participants were asked to learn 64 *kanji* over six weeks. A *kanji* test, consisting of two reading and two writing components, was conducted at the end of the sixth week. On all the tests, international students scored higher than did the local students. Results from the questionnaire indicated that Chinese students had the most knowledge in writing *kanji*, followed by the Korean students, and then the local and Malaysian students. That BS learners outperformed NBS learners was also evident in Machida’s study, where Chinese-background learners scored higher in *kanji* comprehension. This is not surprising given that many Chinese characters are the same as or similar to *kanji*.

These studies by Aoki and Machida show that students’ backgrounds influence how students perform in reading and writing *kanji*. Further to this, Iwashita (2012) investigated cross-linguistic influence as a factor in the writing production of Year 10 learners of Japanese across Australian schools. One-hundred-and-nine written scripts were used (collected from a larger study [Scarino et al., 2011]). They were analysed quantitatively and qualitatively. Out of these samples, 36.4 per cent of the students were BS learners. Six different criteria were used for analysis: content, vocabulary, form and
structure, discourse, use of scripts, and combination. The quantitative results revealed that BS learners outperformed NBS learners in all criteria. Features that clearly distinguished between the two groups were content and form, and structures—BS learners produced richer content using a variety of forms and structures compared to the NBS learners. Qualitative analysis revealed the salient features that could not be identified quantitatively. For example, compared to the NBS learners, BS learners were more familiar with Japanese culture and thus were more able to provide information on this in their writing. Hence, Iwashita (2012) argued that familiarity with aspects of TL culture could enhance L2 performance.

Iwashita and Sekiguchi (2009) examined Japanese L2 writing development in the context of second-year university students in Australia, which is particularly relevant to this study. They investigated whether learners’ backgrounds and modes of instruction (i.e., whether they had studied the language at school or in an intensive first-year course at university) had an influence on their writing development. Thirty-four intermediate level students participated in their study and this group was divided into three different categories:

**Group A:** commenced Japanese studies at university and had a character-based background \((n = 15)\)

**Group B:** studied Japanese at secondary school (post-secondary) and had a character-based background \((n = 5)\)

**Group C:** had a post-secondary, non-character-based background \((n = 14)\).

Two different writing tasks (on different topics) of approximately 300 to 400 characters in length were given to the participants: one in Semester 1 and another in Semester 2. Participants’ composition data were then analysed by examining the structure, sentence production (number of words and sentence complexity), vocabulary richness (word type and word token) and appropriate use of *kanji*. In general, character-
based background students performed better than did others (especially in the use of *kanji*), which indicated that L1 background influenced the learners’ performance. Although the time spent studying Japanese was longer for students with non-character-based backgrounds (Group C), character-based background students who commenced their Japanese study at university (Group A) performed better in the use of *kanji*. This suggested that with *kanji* performance, the L1 background factor had a stronger influence than the length of study. In terms of sentence production, Group B performed slightly worse than the other two groups. It was interesting to note that although Groups A and C had different backgrounds, both groups wrote a similar number of sentences and a similar result was evident in their performance on simple and complex sentences. Across the three groups, similar patterns in the results were seen for word type and word token analysis. Sentence structure was another aspect in which the study found no significant difference among the three groups. Based on their findings, Iwashita and Sekiguchi (2009) suggested that apart from *kanji* use, the learners’ backgrounds did not have a negative impact on their writing development.

In general, BS learners have an advantage over NBS learners (Aoki, 2010; Iwashita & Liem, 2005; Iwashita & Sekiguchi, 2009; Kim, 2012; Scarino, et al., 2011). However, BS learners may not necessarily perform better than did NBS learners if the BS learners are second-or third-generation immigrants (Elder, 1997). In addition, when examining BS learners, exposure to the TL language may not mean they will perform well, unless they make the effort to practice frequently. It is also important to note that when examining *kanji* performance, L1 background has a stronger influence than the length of study.

### 2.6 Gaps in the Research

The research to date has mainly examined L2 writing development in an ESL/EFL context. Some scholars have investigated L2 writing development through
cross-sectional studies (e.g., Cumming et al., 2006; Knoch, Macqueen et al., 2014; Verspoor, Schmid & Xu, 2012), while others have used longitudinal studies (e.g., Knoch, Rouhshad et al., 2014; Storch & Tapper, 2009; Sasaki 2007, 2009, 2011). Nevertheless, little attention has been given to writing development, especially in character-based script languages such as Chinese and Japanese. Although some researchers have examined L2 Japanese writing processes and products (e.g., Pennington & So, 1993) and others have looked at L2 Japanese writing performance (e.g., Iwashita & Sekiguchi, 2009), the amount of research on L2 Japanese writing development remains minimal (e.g., Ishige, 2007; Tanaka 2009) compared to L2 writing development in an ESL/EFL context. Research on writing development other than in the ESL/EFL context is important, as the writing development of a character-based language may differ to that of non-character-based languages (e.g., in accuracy and complexity).

Very few studies have investigated L2 writing development using both a cross-sectional and longitudinal design in a single study. This is because to assess L2 writing development through both cross-sectional and longitudinal studies is very time consuming. Vyatkina’s (2012) research is one of the few to do this, examining L2 German writing to determine writing development for both groups (cross-sectional study) and individuals (longitudinal study). Nevertheless, only two individuals participated in Vyatkina’s research, and this is not a sufficient number to enable comparisons between the cross-sectional studies.

Various factors affect L2 writing development (e.g., instruction, feedback, motivation and background). It is vital that research not only traces writing development, but also examines what leads to that development or the lack thereof. Employing mixed-methods research (i.e., both quantitative and qualitative research) allows researchers to explore writing development in more depth compared to one
method alone. By incorporating qualitative methods, such as interviews, hidden areas of writing development (such as students’ ability to use resources and engage with feedback) can be revealed. For this reason, many researchers (Knoch, Rouhshad et al., 2014; Manchón & Roca de Larios, 2011; Storch & Tapper, 2009; Xudong et al., 2010) have incorporated a mixed-methods approach. Although many (e.g., Knoch, Rouhshad et al., 2014; Storch & Tapper, 2009; Xudong et al., 2010) use questionnaires as the qualitative research method, this approach does not result in as much detail as do interviews. Other studies using interviews (e.g., Manchón & Roca de Larios, 2011; Storch & Hill, 2008) only had a small number of participants. For example, only 15 students participated in the interviews for Storch and Hill’s (2008) study, with 18 in Manchón and Roca de Larios’s (2011) study.

2.7 The Present Study

To address the gaps in the previous research on L2 Japanese writing development, the current study incorporates both cross-sectional studies and longitudinal designs. Moreover, it employs a mixed-methods approach with a larger number of interview participants. Semi-structured interviews obtain both student and teacher perspectives. Through the semi-structured interviews, it has been possible to explain what factors (e.g., instruction, feedback) lead to writing development or the lack thereof.

The current study addresses three main research questions (RQ): one question uses a cross-sectional study (Study 1) and two questions use a longitudinal study (Study 2). These questions are outlined below.

**Cross-sectional study (Study 1) RQ:**

How does the writing ability of L2 Japanese learners at university differ across different proficiency levels?
Longitudinal study (Study 2) RQ:

1. Does writing ability among university students improve in the short term (one semester) as well as in the long term (one year). If so, which aspects of writing develop? Does a study of longer duration make a significant difference?

2. Is there any difference in L2 writing development between character-based learners and non-character-based learners?

This chapter began with a theoretical discussion on L2 writing development, providing a brief overview of some theories employed in L2 writing research. This discussion highlighted that the noticing hypothesis and the idea of ‘practice’ in skill acquisition theory have informed this study. Studies on L2 writing development were then reviewed. These reviews were followed by a discussion of the factors affecting writing development. Based on gaps in the literature, three main RQs have been introduced. Chapter 3 will discuss the measures considered in this study by reviewing the literature on the measures for assessing writing development.
Chapter 3: Discourse Analytic Measures

This chapter discusses the measures considered in Study 1 to examine L2 Japanese writing development through a cross-sectional approach. It begins with a brief description of aspects of the Japanese writing system that influenced the selection of measures for this study. This is followed by a review of the literature on measures for assessing writing development. The current study draws on the most commonly used measures for assessing writing development: complexity, accuracy and fluency (CAF) (Wolfe-Quintero et al., 1998). However, as these measures have primarily been used in research on students learning ESL, they are not all applicable to this study due to the differences between English and Japanese scripts. Therefore, the chapter also reviews other measures used to examine second language (L2) writing development. The measures deemed most relevant to this study are selected and discussed in detail.

3.1 Features of the Japanese Language

The Japanese writing system is very complicated, as the script is a combination of kanji (Chinese characters), hiragana and katagana (Japanese syllabaries referred to as kana) (Kirwan, 2005). In addition, there is a system of romanisation called romaji, which is not usually used with the other three scripts, but is frequently used in other contexts (e.g., train station names, street signs, product names) and as a pronunciation reference for L2 learners.

Hiragana has 46 basic characters and katagana 46. In order to standardise the Japanese writing system, in 1981 the Japanese Ministry of Education (now called the Ministry of Education, Culture, Sports, Science and Technology) designated 1,945 Jōyō kanji, or kanji for everyday use. The Jōyō kanji were revised in 2010; they now include 2,136 kanji characters (KCs) (CosCom Language Service, 2013). Although over 50,000 kanji exist, Jōyō kanji are used most commonly in everyday communication, such as newspapers and business documents.
When Chinese characters were introduced through the Korean peninsula into Japan in the fourth century AD, the characters were given pronunciations similar to that of their Chinese readings (Hinds, 1983). However, by the eighth century, *kana* characters were also used ideographically in a manner that had no reference to the original Chinese pronunciations (Hinds, 1983). *Kana* refers to *hiragana* and *katagana*; *hiragana* is used for case-marking particles and the inflectional endings of verbals and *katagana* are used for foreign loan words (Hinds, 1983). The Chinese characters, *kanji*, usually have two different pronunciations; *on-* *yomi*, which comes from the original Chinese pronunciation; and *kun-* *yomi*, which comes from the native Japanese pronunciation. An example of the *kanji*, 本 with *on-* *yomi* and *kun-* *yomi* is provided in Table 3.1.

Table 3.1

**Example of Kanji 本**

<table>
<thead>
<tr>
<th>Kanji</th>
<th>On-yomi/Kun-yomi</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>本</td>
<td>ホン (hon): <em>on-</em> <em>yomi</em></td>
<td>日本 (nihon): Japan</td>
</tr>
<tr>
<td>Meaning: book; basis</td>
<td>もと (moto): <em>kun-</em> <em>yomi</em></td>
<td>山本さん (Yamamoto-san) Mr./Ms. Yamamoto</td>
</tr>
</tbody>
</table>

All three systems (*hiragana*, *katagana* and *kanji*) can, and often do, appear in a single sentence. An example of this is given here (*romaji* has also been provided to help non-Japanese readers pronounce the Japanese words).

Example:

私はメルボルンに住んでいます。

*Kanji* ひらがな かたがな

Watashi wa meruborun ni sundeimasu

(I live in Melbourne.)

A summary of the Japanese writing system is shown in Table 3.2.
Table 3.2

Summary of the Japanese Writing System

<table>
<thead>
<tr>
<th>Kanji = kun-yomi and on-yomi</th>
<th>Kana = hiragana + katagana</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kun-yomi</td>
<td>On-yomi</td>
</tr>
<tr>
<td>Japanese pronunciation</td>
<td>Chinese pronunciation</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>2,136 Jōyō kanji</td>
<td>46 hiragana</td>
</tr>
<tr>
<td>Chinese characters</td>
<td>Japanese syllabaries</td>
</tr>
</tbody>
</table>

These features of the Japanese writing system should be considered when examining the writing development of L2 Japanese learners, for whom kanji in particular is challenging (Aoki, 2010; Iwashita & Sekiguchi, 2009). Native speakers of English find kanji difficult compared to speakers of Chinese, as native speakers of English have a phonographic language background, whereas speakers of Chinese have a morphographic language background (Kubota, 2005). The phonographic language writing system represents sounds by individual symbols; for example, a phoneme in the English language (Mori, 1998). In contrast, morphographic languages use logography, where a character or symbol represents a word rather than the sound (e.g., Chinese characters) (Mori, 1998). It is more difficult for phonographic language background learners to recognise kanji when compared to those from morphographic language backgrounds. Kanji requires visual information for decoding, but phonographic background learners require phonographic information, which makes it difficult for them to recognise kanji. Moreover, because kanji originated in China (Hinds, 1993), many characters are identical or similar to Chinese characters. This means that Chinese-background learners already know many kanji and can write the characters even if they are unable to pronounce (i.e., speak) them.

Japanese words can be divided into content words, which have their own lexical meaning, and grammatical words, which act as morphological elements (Yamaguchi,
Grammatical words include those that are equivalent in English to inflections, particles, auxiliary verbs, copula and nominalisation (Yamaguchi, 2007). Grammatical words are expressed only in hiragana. Content words are usually expressed in kanji alone; in kanji and hiragana; or in some cases, hiragana alone. In content words expressed in kanji and hiragana, the hiragana functions as okurigana, a suffix that supplements the content part.

The sentence below (an example from Yamaguchi [2007]), which is broken down in Table 3.3, demonstrates the differences between content and grammatical words.

Example:

漢字習得の難しさは日本語教育が始まって以来一貫して指摘されてきたことです。

*Kanji shuutoku no muzukashisa wa nihongo kyouiku ga hajimatte irai ikkan shite shitekisarete kita koto desu.*

Ever since the beginning of Japanese language education for foreigners, the task of learning kanji has been one of the major challenges for learners.

Table 3.3

*Breakdown of Content Words and Grammatical Words*

<table>
<thead>
<tr>
<th>Content words</th>
<th>Grammatical words</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Kanji</strong></td>
<td>okurigana</td>
</tr>
<tr>
<td>漢字習得</td>
<td>の</td>
</tr>
<tr>
<td><em>Kanji shuutoku</em></td>
<td></td>
</tr>
<tr>
<td>難</td>
<td>かし</td>
</tr>
<tr>
<td><em>muzu</em></td>
<td><em>kashi</em></td>
</tr>
<tr>
<td>日本語</td>
<td>が</td>
</tr>
<tr>
<td><em>nihongo</em></td>
<td></td>
</tr>
<tr>
<td>教育</td>
<td>が</td>
</tr>
<tr>
<td><em>kyouiku</em></td>
<td></td>
</tr>
</tbody>
</table>
### Summary of Japanese Words

<table>
<thead>
<tr>
<th>Content words</th>
<th>Grammatical words</th>
</tr>
</thead>
<tbody>
<tr>
<td>起点</td>
<td>まっ</td>
</tr>
<tr>
<td>haji</td>
<td>ma</td>
</tr>
<tr>
<td>以来</td>
<td>し</td>
</tr>
<tr>
<td>irai</td>
<td>shi</td>
</tr>
<tr>
<td>指摘</td>
<td>され</td>
</tr>
<tr>
<td>shiteki</td>
<td>sare</td>
</tr>
</tbody>
</table>

Source: Yamaguchi (2007, p. 74)

A summary of Japanese words is shown in Table 3.4.

Table 3.4

**Summary of Japanese Words**

<table>
<thead>
<tr>
<th>Content words</th>
<th>Grammatical words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kanji only</td>
<td>Hiragana</td>
</tr>
<tr>
<td>Kanji + hiragana (okurigana)</td>
<td></td>
</tr>
<tr>
<td>Hiragana (only in some cases)</td>
<td></td>
</tr>
</tbody>
</table>

The Japanese features discussed in this section indicate that some language features only exist in the Japanese writing system. The literature on writing development in the Japanese context is scarce and those who have examined writing development have done so mostly in an ESL/EFL context. Various measures applied to examine writing development in an ESL/EFL context, as well as in the Japanese context, will be presented in the next section. Additionally, a discussion regarding whether these are appropriate for investigating L2 Japanese writing will be provided.

### 3.2 Complexity, Accuracy and Fluency

Study 1 of this thesis draws on CAF, the most commonly used measure for assessing writing development (Wolfe-Quintero & Kim, 1998). As this is frequently used in an ESL/EFL context, it was worth investigating whether they were also applicable in a Japanese context. Wolfe-Quintero et al. (1998) interpret these measures as follows:
**Complexity:** As L2 learners become more proficient, they are able to write more grammatically and lexically complex sentences.

**Accuracy:** As L2 learners become more proficient, they are able to write more accurately or produce fewer errors when they write.

**Fluency:** As L2 learners become more proficient, they are able to write more fluently or write more in the same amount of time. (p. 4)

These three characteristics of writing development are assumed to progress simultaneously; in other words, the more proficient a learner becomes, the more fluent, accurate and complex she or he will be in their writing (Wolfe-Quintero et al., 1998). However, some studies (e.g., Casanave, 1994) suggest that individual variability exists, and that one characteristic may progress at the expense of another, such as when a trade-off occurs between accuracy and complexity (Skehan, 2009).

### 3.2.1 Complexity

According to Skehan (2009), complexity should be divided into two separate categories: lexical and syntactic. These categories should be separated when evaluated, as they represent different performance areas. The literature on lexical complexity will be discussed here, followed by an examination of studies on syntactic complexity.

#### 3.2.1.1 Lexical complexity

Lexical complexity is concerned with the richness of a writer’s lexicon. Wolfe-Quintero et al. (1998) state that it means ‘a wide variety of basic and sophisticated words are available and can be accessed quickly, whereas a lack of complexity means that only a narrow range of basic words are available or can be accessed’ (p. 69). Past research indicates that lexical complexity is analysed by examining four components: mean word length, lexical variation, lexical density and lexical sophistication. In English, the longer the mean word length, the more complex the lexical item will be. This is not the case with Japanese; therefore, this measure (mean word length) will not
be applicable. Lexical variation, lexical density and lexical sophistication are discussed in the sections below.

3.2.1.1.1 Lexical variation/kanji variation

Type token ratio (TTR), which refers to the total number of different word types divided by total number of words, is the most commonly used measure of lexical complexity (Knoch, 2009). Critics of this measure note that its sensitivity to text length is a failing (Richards, 1987). A negative relationship is evident between sample size and TTR; as the sample size increases, the ratio decreases (Wolfe-Quientero et al., 1998). This may be why Cumming and Mellow (1996) found no significant differences when using this method to measure word variations between ESL learners in different program levels. Even though Cumming and Mellow’s (1996) study did not indicate significant differences, the results showed that means differed in the expected direction. Some researchers (e.g., Arnaud, 1992) collect equivalent samples from each writer, limiting the set to the smallest sample in an attempt to avoid the negative relationship between sample size and TTR. In order to overcome TTR’s limitations, different scholars have applied other methods: the best alternatives are the root type token ratio (RTTR) and the corrected type token ratio (CTTR) (Park, 2013), which are defined below:

1. RTTR, also referred as Guiraud’s index, is calculated by dividing the number of types (V) by the square root of the number of tokens (N)

2. CTTR is calculated by dividing the number of types (V) by the square root of two times the total number of tokens (N)

A few studies (e.g., Daller, Van Hout & Treffers-Daller, 2003; Park, 2013; Vermeer, 2000) suggest that these two measures (RTTR and CTTR) are more useful than TTR when considering only the first few hundred tokens. This is because the manipulation of token count decreases the impact of declining TTR (Park, 2013).
In the Japanese writing system, where words can be written in hiragana, katana and kanji, kanji is considered difficult for many L2 learners to master (Aoki, 2010; Hatasa, 1989; Iwashita & Sekiguchi, 2009; Kirwan 2005; Mori & Shimizu, 2007). Thus, many past studies (e.g., Aoki, 2010; Machida, 2001) have examined L2 learners’ kanji performance as a sign of language development. Therefore in this study, I have considered complexity measures related to kanji. To date, the only study on Japanese language development to have used TTR, RTTR or CTTR to measure kanji complexity is Kikuchi (2012). Specifically, Kikuchi used TTR to analyse kanji complexity. The measure of kanji complexity is not concerned with the number of kanji present in the text, but with how varied the use of kanji is. In other words, the higher the TTR, the more variety of kanji a learner produces. Using this measure provides information on a learner’s degree of knowledge in kanji. The main aim of this study (Kikuchi, 2012) was to investigate the difference in performance of Year 10 students from different backgrounds. Participants were divided into four different groups:

- Group A: Character-based background learners (overseas-born)
- Group B: Character-based background learners (Australian-born)
- Group C: Non-character-based background learners (parents overseas-born)
- Group D: Non-character-based background learners (Australian-born).

Data for this study were obtained from the SAALE project. Two writing tasks were undertaken as part of the students’ assessment. Writing scripts from 78 participants were collected for this study. In relation to kanji complexity, in general, learners were unable to produce a wide variety of kanji regardless of their background; that is, repetition of the same kanji was evident among all groups.

An example of how to measure kanji complexity is shown below; this is taken from Kikuchi (2012). In the typed version, kanji that are used once are highlighted in yellow and repeated kanji are highlighted in green. Non-highlighted scripts are in
hiragana, katagana or English. This learner only used ten different kanji types: 私，十，五，日，本，行，語，好，一，母. Twenty-five kanji tokens were used; the TTR was $10/26 = 0.38$, which is considered quite low.

Example script for measuring complexity:

私の名前はジョーデイといって十五さいです。りゅうがくせいになって、日本に行きたいたいです。日本に行きたくても日本語をべんきょうすることが好きです。しゅみはまんがをよんで、やきゅうのしあいをみて、じょうばをします。日本に行ったらいいたいけんをしてたくさんあたらしいともだちにあってたのしみにしています。おはん好きおんがくはJ-POPですけれどもうちでJ-POPのCDがありません。それから日本に行ったらたくさんJ-POPのCDをかいます。私の母は日本に行って、日本はすごいをはなしました。

My name is Jodi and I am 15 years old. I want to go to Japan as an exchange student. I don’t go to Japan but I like to study Japanese. My hobby is to read comics, watch baseball games and horse riding. When I go to Japan I would like to have a good experience and make new friends, I am looking forward to that.

The music I like most is J-Pop But I don’t have J-Pop CD at home. Furthermore,
when I go to Japan I will buy many J-Pop CDs. My mother went to Japan and told me that Japan was great.

3.2.1.1.2 Lexical density

Lexical density, as developed by Ure (1971), refers to the proportion of content words relative to the total number of words. This measure is often used to compare the lexical density of spoken and written registers or genres. Past research (e.g., Stubbs, 1996; Ure, 1971) has found that spoken texts have less lexical density compared to written text. Conversely, Park (2013) used this measure to investigate the differences in lexical density in different written discourse modes and found no effect. This indicates that this measure is useful when comparing spoken and written registers but not for examining the differences in lexical density across different written registers. Therefore, lexical density was not considered in my pilot study.

3.2.1.1.3 Lexical sophistication

Wolfe-Quintero et al. (1998) reviewed a number of studies (e.g., Hyltenstam, 1988; Linnarud, 1986) that used lexical sophistication—the total number of sophisticated lexical words divided by total number of lexical words—and found that this measure is related to second language development. In order to calculate lexical sophistication, lexical words on a special word list (such as the Academic Word List [AWL] [Coxhead, 2000]) must be identified in a written sample. For example, Storch and Tapper (2009) examined the vocabulary used in students’ written work against the framework of academic words on the AWL. They counted the number of AWL types and tokens in each student’s writing at two different intervals, nine weeks apart. The result suggested a significant improvement in students’ use of academic vocabulary over time.

Studies related to lexical sophistication in the Japanese language are scarce and no studies on academic vocabulary exist. As the students in this study were neither
learning nor being exposed to academic writing, this measure was not relevant and thus was not considered in my pilot study.

After reviewing the literature on lexical complexity, examining *kanji* complexity was deemed important. The following measures were chosen for my pilot study (they will be categorised as *kanji* complexity under complexity):

1. KC TTR
2. KC RTTR
3. KC CTTR
4. *kanji* word (KW) TTR
5. KW RTTR
6. KW CTTR.

The next section provides a brief discussion of the literature on syntactic complexity and the measures chosen for the pilot study.

3.2.1.2 *Syntactic complexity*

Syntactic complexity is concerned with grammatical complexity; in other words, how varied or sophisticated a learner’s grammatical structures are. It is not concerned with how many production units there are in a written composition, but with how complex they are.

There are a number of ways to measure syntactic complexity. After reviewing the literature, those deemed appropriate for the Japanese language were chosen and are discussed in this section. First, studies using T-units to measure syntactic complexity are briefly discussed. Second, the ‘total frequency of use of certain forms considered to be sophisticated’, identified in Norris and Ortega (2009) is outlined. This measure is important, as one Japanese grammatical feature called ‘te-form’ is often used in a number of ways and is considered sophisticated (a detailed discussion will be provided in the following section). Third, sentence complexity (Iwashita & Sekiguchi, 2009),
clause conjunctions (Tanaka, 1997) and sentence-ending variables (Tanaka, 1999; Sasaki and Kawaguchi) – measures used by other scholars to investigate L2 Japanese writing are outlined.

3.2.1.2.1 Measuring syntactic complexity with T-units

As briefly mentioned in Chapter 2 (Section 2.3.1.2), the T-unit was developed by Hunt in 1965. According to Hunt (1970), a T-unit is defined as ‘the shortest [unit] in[to] which a piece of discourse can be cut without leaving any sentence fragments as residue’ (p. 188). It includes one main clause and all subordinate clauses attached to it (Hunt, 1970, p. 20).

Since the introduction of the T-unit, the T-unit complexity ratio has been widely used in L2 research (Gaies, 1980). Wolfe-Quintero et al. (1998) identified that the T-unit complexity ratio (number of clauses per T-units) was one measure that could distinguish the most between different proficiency levels (in an ESL/EFL context). Accordingly, this section will discuss T-unit complexity in detail.

Knoch (2009) explains that a T-unit complexity ratio of two would mean that ‘on average each T-unit consists of one independent clause plus one other clause’ (p. 87). The T-unit complexity ratio’s purpose is to measure grammar complexity (assuming that more clauses exist per T-unit in complex writing) (Knoch, 2009).

However, past studies have shown mixed results in relation to syntactic complexity and writing proficiency. For example, both Hirano (1991) and Flahive and Snow (1980) found that there was a significant relationship between the T-unit complexity ratio and program level. However, other studies such as Cumming et al. (2005) and Banerjee et al. (2007) found no significant results. Banerjee et al.’s (2007) study on IELTS essays found that similar numbers of clauses per T-unit existed across proficiency levels. Wolfe-Quintero et al. (1998) noted that this measure related to the program level or rating scale and this might explain the mixed findings.
In relation to the Japanese language, although no research has been conducted using T-units to analyse writing complexity, several scholars have attempted to analyse speaking performance using T-units (e.g., Harrington, 1986; Iwashita, 2006; Taguchi, 2008; Tamaru, Yoshioka & Kimura, 1993). These studies provide some mixed findings about the usefulness of T-units. For example, Harrington (1986) examined the oral performances of 14 learners of L2 Japanese to ascertain if T-unit measures were reliable. Some degree of discrimination was found between learners with the average T-unit length and the average length of an error-free T-unit. Harrington (1986) concluded that T-unit measures possessed limited usefulness for oral performances in Japanese.

However, Iwashita’s (2006) study obtained different results. Iwashita examined the effectiveness of a number of syntactic complexity measures (including various T-unit measures) to determine which one effectively predicted oral proficiency. Thirty-three learners at two levels of proficiency (low and high) participated. The complexity measures investigated are shown in Table 3.5.

Table 3.5

<table>
<thead>
<tr>
<th>Syntactic complexity category</th>
<th>Syntactic complexity measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Production unit</td>
<td>T-unit, clause, verb phrase, word</td>
</tr>
<tr>
<td>Length</td>
<td>No. of words, no. of words types, no. of words per T-unit, no. of words per clause</td>
</tr>
<tr>
<td>General complexity measure</td>
<td>No. of clauses per T-unit</td>
</tr>
<tr>
<td>Coordination measures</td>
<td>No. of ICs per T-unit, no. of ICs per clause</td>
</tr>
<tr>
<td>Subordination measures</td>
<td>No. of DCs per T-unit, no. of DCs per clause, no. of VPs per T-unit, no. of VP per clause</td>
</tr>
</tbody>
</table>

Note: IC = independent clause; DC = dependent clause; VP = verb phrase

In contrast to Harrington’s (1986) study, Iwashita (2006) concluded that the length of T-units and the number of clauses per T-unit were the best predictors of syntactic complexity. Although Iwashita’s (2006) study investigated oral proficiency, it also examined the Japanese language; the two measures (length of T-unit and clauses
per T-unit) were effective, these measures were thus worth considering in my pilot study.

Tanaka (1997) did not use the T-unit in his study, instead focusing on words and clauses per sentence to measure syntactic complexity. This study examined the writing development of L2 learners of Japanese before and after in-country experience. Both measures proved useful, as they revealed developmental changes.

3.2.1.2.2 Total frequency of use of certain forms considered sophisticated: Te-form.

Te-form ending verbs are considered grammatically complex forms in Japanese (Taguchi, 2008). The te-form verbs are used for various functions (Makino & Tsutsui, 2004; Taguchi, 2008):

- they link two actions; for example, arui-te kaeru (I walk home)
- they combine with the auxiliary iru (exist) to indicate continuous action; for example, tabete iru (eating)
- they combine with the auxiliary kudasai (please) to show polite commands; for example, tabete kudasai (please eat)
- they combine with the auxiliary moiidesu (you may) to express permission; for example, tabe-te moiidesu (you may eat).

Taguchi (2008) examined chunk learning to reveal the developmental changes in L2 Japanese speaking. Of the various factors she investigated, the results indicated that chunk learning allowed even beginners to use the complex te-form at an early stage. In another study conducted by Kato and Kobayashi (1985), the numbers of te-form verbs were counted to see which were used most frequently after three months of instruction in beginner learners. Their results suggest that te-iru (‘doing’) was the most frequently used te-form. Both studies (i.e., Taguchi, 2008 and Kato & Kobayashi, 1985) suggest that looking at a te-form is useful for examining complexity. Simply counting the
number of te-form may not provide accurate results; hence, in my pilot study the ratio of te-form per words was considered as a measure.

3.2.1.2.3 Sentence complexity

Iwashita and Sekiguchi (2009) developed a different measure of syntactic complexity. They categorised every sentence in their Japanese L2 learners’ compositions as either simple or complex in order to evaluate the learners’ development. According to Iwashita and Sekiguchi (2009, p. 9), simple sentences contain only one verb and one main idea.

Example:

オーストラリア人はリラックスしすぎだ。

*Oosutoraria jin wa rirakkusu shisugida.*

* Australians are too relaxed.*

In contrast, complex sentences contain two or more verbs, consist of two or more clauses and contain more than one main idea (Iwashita & Sekiguchi, 2009, p. 10).

Example:

雪が降ると電車がとまる。

*Yuki ga furu to densha ga tomaru.*

*If snow falls, trains stop.*

Participants were divided into three groups, which allowed the researchers to factor in their L1 background and L2 study experience:

**Group A:** learners who had commenced their Japanese studies at university and had a character-based language background (*n* = 15)

**Group B:** post-secondary learners who had studied Japanese at secondary school and had a character-based language background (*n* = 5)

**Group C:** learners at the post-secondary stage who had a non-character-based language background (*n* = 14)
Interestingly, Group A and C showed similar levels of performance in terms of simple and complex sentences, while Group B performed slightly worse compared to these two groups. The mixed results from this study may have been due to the way in which the researchers measured the participants’ performance (i.e., in how the participants were grouped). Nevertheless, Iwashita and Sekiguchi’s (2009) relatively small sample size creates difficulties in making definitive conclusions. Despite this, it was still worth considering this measure for my pilot study.

3.2.1.2.4 Clause conjunctions

Tanaka (1997) examined two types of conjunctions: general and clause conjunctions (which he did not categorise under a specific category). In this study, general conjunctions are categorised into ‘cohesion’ (this will be discussed in Section 3.2.4.2.1) and clause conjunctions are categorised into ‘syntactic complexity’. In a clause conjunction, words such as conjunctive particles kara (reason) are used to combine two sentences into a single sentence. Clause conjunctions, therefore, join a dependent clause with an independent clause to make a complex sentence (Tanaka, 1997).

Example:
今日は朝早くおきたから眠い。

*Kyou wa asa hayaku okitakara nemui.*

Because I woke up early today, I am sleepy.

In his study, Tanaka (1997) considered six dependent clause features for clause conjunctions; these are as follows:

1. conjunctive particles such as *ga* (but), *node* (because)
2. conjugational forms such as the *reba-form*, the *tari-form*
3. relativisation forms such as those that create relative clauses
4. nominalisers such as *koto*, *wake*, *tokoro*
5. quotation particles such as to (that), yo ni (as)

6. nominal endings such as da/dearu (is).

By counting the occurrences of clause conjunctions, Tanaka (1997) could see the developmental changes in learners’ writing before and after in-country experience. However (as mentioned earlier in the discussion of te-form), simply counting the words may not give accurate results. Therefore, the ratio of clause conjunctions to words was considered for the pilot study.

3.2.1.2.5 Sentence-ending variables

Sentence-ending variables are considered important when examining the ‘naturalness’ of writing in Japanese (Tanaka, 1997), and in this study they are categorised into syntactic complexity. Sasaki and Kawaguchi (1994) compared the use of sentence-ending forms between native and non-native speakers of Japanese. The authors found that, compared to native speakers, non-native speakers used more propositional (e.g., suru meaning ‘to do’) and explanatory endings (e.g., suru no da meaning ‘do’), as well as fewer epistemic endings (e.g., suru daroo meaning ‘probably do’). The findings also revealed that as native speakers mature, they tend to use less propositional endings. In relation to epistemic forms, Sasaki and Kawaguchi’s (1994) study found that non-native speakers and Japanese junior high school students employed to omou (‘think that’) and ki ga suru (‘feel that’) frequently. Interestingly, Japanese university students used different expressions; for example, daro ka (‘wonder if’) and kamoshirenai (‘might’).

Tanaka (1997) compared the use of sentence-ending variables between pre- and post-Japanese visit groups. He examined the sentence-ending variables from the perspective of modality, and categorised them into eight types (see Table 3.6) (Tanaka 1997, p. 70).
Table 3.6

Eight Types of Sentence-ending Variables

<table>
<thead>
<tr>
<th>Sentence-ending variables</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Propositional endings</td>
<td>suru (do)</td>
</tr>
<tr>
<td>Negative endings</td>
<td>shinai (do not), dewanai (is not)</td>
</tr>
<tr>
<td>Past endings</td>
<td>shita (did), datta (was)</td>
</tr>
<tr>
<td>Explanatory endings</td>
<td>suru no da (do), suru wake da (so it means)</td>
</tr>
<tr>
<td>Judgement endings</td>
<td>suru beki da (should do)</td>
</tr>
<tr>
<td></td>
<td>shita ho ga yoi (had better do)</td>
</tr>
<tr>
<td>Epistemic endings</td>
<td>to omou (think that), suru kamoshirenai (might do)</td>
</tr>
<tr>
<td>Expressive endings</td>
<td>shitai (want to do), shiyo (let us do)</td>
</tr>
<tr>
<td>Nominal endings</td>
<td>da/dearu (is)</td>
</tr>
</tbody>
</table>

Tanaka (1997) found that propositional endings were used most frequently. Nevertheless, the occurrence percentage of propositional endings declined after students’ in-country experience. In terms of propositional endings, the findings from both Tanka (1997) and Sasaki and Kawaguchi (1994) suggest that as a learner’s proficiency increases, the use of propositional endings declines. Further, a significant increase was evident in the occurrence of expressive endings, past endings and epistemic endings with the post-Japanese visit group.

Both Sasaki and Kawaguchi (1994) and Tanaka (1997) concluded that examining sentencing-ending forms allowed them to see a certain pattern of development in learners’ writing. As such, sentence-ending forms were chosen for my pilot study and the ratios of sentence-ending forms to words were considered as a measure.

After reviewing the literature on syntactic complexity, the following measures were chosen for my pilot study:

1. number of words per T-unit
2. number of clauses per T-unit
3. words per sentence
4. clauses per sentence
5. ratio of total frequency of te-form to words
6. sentence complexity (ratio of simple and complex sentences to sentences)
7. ratio of clause conjunctions to words
8. ratio of sentence-ending forms to words.

3.2.2 Accuracy

Foster and Skehan (1996) define accuracy as ‘freedom from error’. This can be measured by examining L2 learners’ written production and determining if the L2 learner demonstrates target-like use. The current study considers error as ‘a linguistic form or combination of forms which, in the same context and under similar conditions of production, would, in all likelihood, not be produced by the speakers’ native-speaker counterparts’ (Lennon, 1991, p. 182).

Various types of accuracy measures are discussed in the literature. Polio (1997) examined linguistic accuracy measures, focusing on ESL writing, and concluded that three main types of measures existed:

1. holistic ratings of accuracy
2. error-free units
3. error-frequency measures.

These are briefly discussed in the following section.

3.2.2.1 Holistic ratings of accuracy

Holistic ratings of accuracy are often used as one component among others in a composition rating scale (Polio, 1997). Although Polio (1997) indicated this method enables researchers to evaluate accuracy quickly compared to the other two measures (i.e., error-free units and error-frequency measures), she suggests that the scales often use ambiguous terms. For example, some scales use words such as ‘frequent’ and
‘occasional’ to quantify the number of errors, and others use ‘significant’, ‘effective’ or ‘sophisticated’ to characterise the quality of the language (Polio, 1997).

Despite Polio’s (1997) concerns regarding the ambiguity of holistic ratings, recent studies (e.g., Cumming et al., 2006; Gebril, 2009) suggest that significant differences can be seen using such ratings to compare high- and low-proficiency level learners. For example, Cumming et al. (2006) investigated learners’ grammatical accuracy using a 3-point scale akin to that used in Hamp-Lyons and Henning (1991):

1. many severe errors, often affecting comprehensibility
2. some errors, but comprehensible to a reader
3. few errors, and comprehensibility seldom obscured for a reader.

3.2.2.1.2 Error-free units

In order to analyse the accuracy of an L2 ESL/EFL writer, many researchers count EFT and/or error-free clauses (EFC) as an objective way of measuring accuracy. However, this type of measure can be subjective, as researchers must decide what constitutes an error: this may vary from researcher to researcher (Wolfe-Quintero et al., 1998). Moreover, these measures may not distinguish between proficiency levels as the same weight is given to one, or minor errors as units, with many or very grave errors.

Robb, Ross and Shortreed (1986), Casanave (1994) and Ishikawa (1995) have used this method to examine accuracy. After reviewing some of the studies on EFT and EFC, Polio (1997) concluded that EFTs and EFCs enable researchers to obtain information on the quantity, but not the quality, of errors.

Most scholars (e.g., Kato & Kobayashi, 1985; Kubota, 2005) who have examined the accuracy of L2 Japanese learners have counted errors (i.e., the error-frequency measures described below), whereas Tanaka (1997) used the ratio of error-free sentences to sentences, and words per error-free sentence as accuracy measures. Words per error-free sentence showed a slight increase in the mean number between
pre-Japan visit learners and post-Japan visit learners, whereas the ratio of error-free sentences to sentences decreased slightly. The small number of participants ($N = 10$) in Tanaka’s study may not enable a definite conclusion to be made regarding the usefulness of these measures. Despite this, these two measures were still considered in the pilot study.

3.2.2.1.3 Error-frequency measures

Some scholars (e.g., Bitchener & Knoch, 2008; Kepner, 1991) have counted the number of errors or calculated accuracy as a percentage of correct usage as opposed to the number of EFTs or EFCs. An example of how the percentage of correct usage is calculated is as follows: in any one script, three correct uses of the targeted features (e.g., ‘a’ and ‘the’) from ten obligatory occasions means a 30 per cent accuracy rate.

With Japanese writing, this measure of obligatory context can be applied to the correct usage for particles, *kanji*, and *kana*. Research indicates that L2 learners of Japanese also struggle with how to use particles. For example, Kato and Kobayashi (1985) counted the number of times learners used double particles (see Section 3.2.3.1 for further details on fluency). Among other measurements, the researchers found that beginner L2 Japanese learners struggled with double particles and in particular, tended to overuse the double particle *niwa* and *dewa*, barely using other double particles apart from these. Kato and Kobayashi (1985) did not indicate the number of errors made by the learners, but suggested that the overuse of these double particles often led to errors. Thus, investigating the correct usage of particles was deemed important and considered in the pilot study. However, instead of simply counting the particles, the percentage of particles in the written script was considered.

Although Kubota’s (2005) study aimed to investigate spelling correction strategies employed by intermediate learners of Japanese, Kubota’s classification is useful for measuring accuracy. It gives a general idea of what kind of errors learners of Japanese L2 may make. The errors identified in Kubota (2005) relevant to this study are presented below. Table 3.7 shows the types of kanji errors and Table 3.8 shows possible types of kana errors made by Kubota’s participants.

Table 3.7

Types of Kanji Errors

<table>
<thead>
<tr>
<th>Types of kanji errors</th>
<th>Example of an error</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. error</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confusion with morphographically similar kanji</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. 恋になりました（恋：the kanji on the bottom is 心 whereas the correct kanji should be 変）</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. 変になりました</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A misshapen component of the kanji</td>
<td>1. 建物</td>
<td>When learners handwrite kanji, it is sometimes poorly written</td>
</tr>
<tr>
<td>2. 建物</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problems with okurigana</td>
<td>1. 入る</td>
<td>When a word is expressed with kanji and hiragana, hiragana functions as okurigana, an ending that supplements the word. In this case, there is an additional hiragana in the word, 入る instead of 入る</td>
</tr>
<tr>
<td>2. 入る</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transfer from Chinese characters to Japanese</td>
<td>1. 伝統的</td>
<td>Background learners sometimes get confused with the Chinese characters and kanji</td>
</tr>
<tr>
<td>2. 伝統的</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homonyms</td>
<td>1. 始めて</td>
<td>始めて and 初めて are both pronounced ‘hajimete’. 始めて means ‘to begin’ but 初めて means ‘first time’.</td>
</tr>
<tr>
<td>2. 初めて</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The absence of one kanji from a compound KW</td>
<td>1. 小校</td>
<td>The kanji 学 is missing in the 小校.</td>
</tr>
<tr>
<td>2. 小学校</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3.8

*Types of Kana Errors*

<table>
<thead>
<tr>
<th>Types of kana errors</th>
<th>Example of an error</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. error</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. correct version</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A misshapen kana</td>
<td>1. 变わりました</td>
<td>When learners handwrite kanji, it is sometimes poorly written</td>
</tr>
<tr>
<td></td>
<td>2. 変わりました</td>
<td></td>
</tr>
<tr>
<td>A missing component</td>
<td>1. なりました</td>
<td>The hiragana す is missing at the end.</td>
</tr>
<tr>
<td></td>
<td>2. なります</td>
<td></td>
</tr>
<tr>
<td>A wrong kana</td>
<td>1. 思し</td>
<td>Hiragana し is used instead of い.</td>
</tr>
<tr>
<td></td>
<td>2. 思い</td>
<td></td>
</tr>
<tr>
<td>Transfer from Chinese</td>
<td>1. 傳統的</td>
<td>Background learners sometimes get confused with the Chinese characters and kanji</td>
</tr>
<tr>
<td>characters to Japanese</td>
<td>2. 伝統的</td>
<td></td>
</tr>
<tr>
<td>Writing a loan word</td>
<td>1. ヴァラエティ</td>
<td>When the loan word ‘variety’ is written as ヴァラエティー, the ヴァ will be close to the English ‘va’, but the ‘Japanised’ version is written with バ with the ‘ba’ pronunciation, hence ‘baraety’ in romaji</td>
</tr>
<tr>
<td>approximating the original word rather than using the ‘Japanised’ version</td>
<td>2. バラエティー</td>
<td></td>
</tr>
</tbody>
</table>

Polio (2001) notes most studies do not include spelling errors when considering accuracy. However, Bereiter (1980) does suggest that measurements such as spelling and punctuation must be considered when assessing accuracy, as writing is different to speech. Thus, examining the correct usage of kanji and kana was also considered for the pilot study and Kubota’s error classification was used as a base when looking at errors.

After reviewing the literature on accuracy, the following measures were chosen for my pilot study:
1. words per error-free sentence
2. ratio of error-free sentences
3. ratio of EFCs
4. ratio of error-free KCs
5. ratio of error-free KWs
6. ratio of error-free kana
7. ratio of error-free particles.

In the following section, studies that used fluency as one measure are discussed and the measures selected for the pilot study are identified.

3.2.3 Fluency

Fluency measures indicate how fluently an L2 writer can produce written material. Skehan (2003) identified four measures in speaking fluency that can also be related to writing fluency (Abdel Latif, 2013; Knoch, 2009):

1. **Breakdown fluency:** in terms of speech, this is measured by silence, whereas in writing it is measured by a break or pause in the writing process.

2. **Repair fluency:** in terms of speech, this refers to reformulations, replacements, false starts and repetition. In writing, it can be measured by the number of revisions a learner goes through while composing a writing task.

3. **Speech/writing rate (temporal aspect of fluency):** an example of this is the number of syllables per minute or words per minute.

4. **Automatisation:** this is the length of bursts that occur between breaks.

Breakdown fluency and automatisation cannot be deduced by only examining the writing product. Only transcripts of think-aloud protocols, videotapes of writing sessions or keystroke software will allow researchers to examine these features (see Chenoweth & Hayes, 2001; Lindgren, 2005; Wengelin et al., 2009). According to
Knoch (2009), ‘only repair fluency and temporal aspects of writing (writing rate) can be measured on the basis of a writing product’ (p. 85). Repair fluency would not be appropriate for my study, as the written product collected could not record the number of revisions a student had made. Therefore, in my study different types of temporal aspect were considered. These are discussed in detail below.

3.2.3.1 Measures of temporal aspects of fluency

According to Wolfe-Quintero et al. (1998), ‘the primary way to measure fluency is count the number, length, or rate of production units’ (p. 14). Sentences, clauses, T-units and phrases are considered production units (Wolfe-Quintero et al., 1998). Studies using the number of words, number of clauses, number of sentences and number of kanji to measure fluency are discussed in the following sections.

3.2.3.1.1 Number of words

To examine the temporal aspects of writing rate fluency in an ESL context, many researchers (e.g., Cumming et al., 2005; Hirano, 1991; Ishikawa, 1995; Kennedy & Thorp, 2002) have applied the number of words produced for writing composed under test conditions (with time limitations) as their measure. Although Wolfe-Quintero et al. (1998) indicated that counting the number of words is a primary way to measure fluency, after reviewing various literature, they concluded it was not a reliable measure of development, due to the mixed results. Even studies conducted after Wolfe-Quintero et al.’s (1998) publication do not indicate whether counting the number of words is an effective measure. For example, Kennedy and Thorp (2002) examined the differences in writing performance at three different IELTS levels (Levels 4, 6 and 8) and found a difference between essays at three levels, with Level 4 writers producing the shortest essays. However, a considerable overlap was evident between the minimum and maximum number of words in compositions at the three levels. Similarly, Cumming et al. (2005) found differences in writing performance in TOEFL essays at Levels 3 and 4,
and between essays at Levels 3 and 5. Nevertheless, no statistically significant
differences existed between Levels 4 and 5, with descriptive statistics indicating a small
increase in the number of words used.

Compared to studies in English, few studies on fluency in Japanese exist. Those
that have examined fluency using the number of indicate this is an effective measure.
Yamamoto (2007) conducted a study on the writing performances of L1 Japanese
tertiary level learners (i.e., native Japanese speakers studying Japanese). Thirty students
participated in her study. After the participants took a Japanese proficiency test for L1
learners of Japanese, 20 were allocated to Group A (standard of Japanese writing at
Year 10 level or below) and ten were allocated to Group B (standard of Japanese
writing at Year 12 level). Word length was used as a measure to examine fluency.
Group B outperformed Group A. In other words, Group B (students with a higher level
of Japanese writing skill) produced more words than did Group A (students with lower
levels of Japanese writing skills). This measure was a useful indicator of proficiency, as
it revealed clear differences between Group A and B (Yamamoto, 2007).

Kato and Kobayashi’s (1985) study examined the writing of beginner L2
learners of Japanese after an intensive three-month instructional period. Kato and
more words showed developmental progress, whereas those students who wrote fewer
words showed little developmental progress.

Similarly, Tanaka (1997) used the number of words to examine the development
of Japanese writing skills through in-country experience. The results suggested that the
number of words was a useful measure, as it showed clear differences between pre-
Japan and post-Japan learners. The same measure of fluency (i.e., counting the number
of words) was applied in a study I conducted (Kikuchi, 2012). My study also indicated
that higher-proficiency students produced more words in a given amount of time. Counting the number of words as a measure of fluency is therefore a useful gauge of language proficiency and development. Hence, this measure was applied in my pilot study.

3.2.3.1.2 Number of clauses

Other than counting the number of words, several other measures can determine fluency, such as counting the number of clauses and sentences (Wolf-Quintero et al., 1998). Very few researchers have employed the number of clauses in research on English or Japanese to measure fluency. Those researchers (e.g., Ishikawa, 1995; Kameen, 1979) who have used the number of clauses to examine fluency in an ESL context found no significant effect in proficiency. For example, Ishikawa (1995) examined beginner-level students’ writing at the start and after a three-month instruction. Kameen (1979) compared well-written and poorly written essays. In contrast, those researchers (Tanaka, 1997; Tashiro, 2007) who used the number of clauses to examine writing development of L2 Japanese found significant differences. As such, this measure is indicated as useful for examining fluency. Tanaka’s (1997) study investigated students’ writing before they went to Japan and after one year of in-country experience in Japan. Here, the number of clauses in their writing had increased after spending time in Japan.

3.2.3.1.3 Number of sentences

The literature reviewed by Wolf-Quintero et al. (1997) indicated that all six studies (Homburg, 1984; Ishikawa, 1995: Group 1 and Group 2; Kameen, 1979; Kawata, 1992; Perkins, 1980) that used the number of sentences to measure fluency found no significant effect. Wolf-Quintero et al. (1997) suggested that this may be because low-proficiency learners can produce simple sentences, while high-proficiency can produce more complex sentences.
Regarding research in Japanese, Tanaka (1997) and Tashiro (2007) both used the number of sentences to measure fluency. Tanaka (1997) identified developmental changes in learners before and after in-country experience in Japan, and Tashiro (2007) noted differences in proficiency between learners using this measure.

3.2.3.1.4 Number of kanji

As kanji in particular remain challenging for L2 learners (Aoki, 2010; Iwashita & Sekiguchi, 2009; Mori & Shimizu, 2007), using the number of kanji to measure fluency is an effective measure to distinguish between different levels. This is evident in both Iwashita and Sekiguchi (2009) and Kikuchi’s (2012) studies. However, these studies examined hand-written texts and it is important to note that this measure may not be effective when examining a typed text. This is because the kanji are generated by the computer, not the learner (the learner only needs to type the romanised version to type the kanji).

Similar to Kikuchi (2012), in this current study, kanji will be divided into KC and KW. The reason for this distinction is that more than one KC can make up a single KW. For example, the word for ‘Kyoto’ is made up of two KCs, 京 and 都, but one KW 京都 (see Figure 3.1).

<table>
<thead>
<tr>
<th>KC</th>
<th>KW</th>
</tr>
</thead>
<tbody>
<tr>
<td>京</td>
<td>京都</td>
</tr>
<tr>
<td>都</td>
<td>1 KW</td>
</tr>
</tbody>
</table>

2 KC

Figure 3.1. Example of the word ‘Kyoto’ (京都).

Based on the above literature, the following measures of fluency have been chosen for my pilot study:

1. number of words (overall)
2. number of clauses
3. number of sentences
4. number of KC
5. number of KW.

3.3 Other Measures that Examine Writing Development

Although CAF has been widely used as a measure for writing development, other measures, including mechanics, cohesion and coherence, are also appropriate. These will be discussed in this section. Studies conducted using these measures focus mainly on students learning ESL. Some measures only applicable to Japanese—such as *desu/masu*—will also be discussed.

3.3.1 Mechanics

In English, mechanics include paragraphing, spelling, punctuation, capitalisation and indentation (Knoch, 2009; Polio, 2001). According to Polio (2001), very few studies have attempted to quantify these aspects. Different types of spelling errors have been included in the accuracy measures for this thesis (see Section 3.2.2) and therefore will not be included as part of the mechanics section. Punctuation is not used as often in Japanese as it is in English; in addition, Japanese does not use capitalisation, and therefore this was not included in my pilot study.

Kennedy and Thorp (2002) examined the number of paragraphs produced by candidates at Levels 4, 6 and 8 in IELTS tests and found that 10 per cent of the writers at Level 4 produced one paragraph and those at Level 6 produced four or more paragraphs. Paragraphing was not be considered for my pilot study, as I believe that counting the number of paragraphs in a written composition is not a sufficient indicator of writing development; some learners may write many short paragraphs and others may write one or two paragraphs with many sentences.

3.3.2 Cohesion and coherence

According to Spiegel and Fitzgerald (1990), cohesion is defined as ‘the linguistic features which help to make a sequence of sentences in a text i.e. give it texture’ (p. 49). Coherence is referred as the interconnection of the cohesive parts
(Halliday & Hasan, 1976). Further, discourse is created by the two elements (i.e., cohesion and coherence), which indicates a relationship exists between discourse, cohesion and coherence. The next section introduces the literature on cohesion, followed by studies on coherence.

### 3.3.2.1 Cohesion

Halliday and Hasan (1976) suggest it is possible to achieve communication without cohesion; however, having a cohesive text helps guide the reader to understand the text units better. Halliday and Hasan (1976) have identified five types of cohesive devices, which are indicated below:

1. **Reference**: this is ‘the relation between an element of the text and something else by reference to which it is interpreted in the given instance’ (Halliday & Hasan, 1976, p. 308) and is divided into personal, demonstrative and comparative.
2. **Substitution**: the use of terms such as ‘the’ (for nouns), ‘do so’ (for verbs) ‘so’ (for clauses) to replace a word or string of words.
3. **Ellipsis**: the omission of an item.
4. **Conjunction**: this can be additive, adversative, causal and temporal.
5. **Lexical cohesion**: the repetition of a noun phrase and is divided into two types: reiteration and collocation.

Several studies (e.g., Neuner, 1987; Tangkiengsirisin, 2010; Faigley & Witte, 1981) have applied Halliday and Hasan’s (1976) categories of cohesion to examine writing performance. Faigley and Witte (1981) examined college-level L1 English essays in the US and compared the cohesion of high-level and low-level written compositions. They found that a higher density of cohesive ties existed in high-level essays. Approximately 30 per cent of all words contributed to cohesion ties in high-level essays, while only just over 20 per cent of the words contributed to cohesion ties.
in low-level essays. Moreover, most of the lexical ties in low-level essays were repetitions, whereas high-level essays depended more on lexical collocation. In a more recent study, Tangkiengsirisin (2010) used a pre- and post-test design to determine if improvements in cohesion were evident in the written compositions of Thai postgraduate students \((N = 60)\). Participants were divided into an experimental group \((n = 30)\) and a control group. The experimental group received feedback on their written assignments, focusing specifically on the use of cohesion. A significant improvement of cohesion in the writing of the experiment group was apparent. Tangkiengsirisin (2010) suggests that although cohesion is a useful linguistic element that indicates well-connected writing, it may not measure the overall quality of writing appropriately.

Several scholars (Kato & Kobayashi, 1985; Tanaka, 1997; Tashiro, 2007) have examined conjunctions to investigate cohesion. These studies indicate that counting conjunctions is useful for determining learners’ developmental changes.

For example, Tanaka (1997) examined two types of conjunctions—general and clause (the latter is categorised under syntactic complexity in my study; see Section 3.2.1.3.3). General conjunctions in Tanaka’s (1997) study refer to conjunctions such as *dakara* (therefore) to relate two independent sentences without making a complex sentence.

Example:

今日は朝早くおきた。だから眠い。

*Kyou wa asa hayaku okita. Dakara nemui.*

I woke up early today. Therefore, I am sleepy.

Tanaka’s (1997) study revealed developmental changes of general conjunction use in learners’ writing before and after in-country experience.

After reviewing the literature, the measure of general conjunction was deemed appropriate and was considered in my pilot study. However, as opposed to many
scholars (Kato & Kobayashi, 1985; Tanaka, 1997; Tashiro, 2007) who simply count the amount of cohesion in a written text, the current study used the ratio of general conjunctions to words.

3.3.2.2 Coherence

No consensus on the definition of coherence exists, as various scholars understand it differently (for one definition, see Grabe & Kaplan, 1996). From the interpretations of various scholars (e.g., Chiang, 1999; Lee, 2002), it would be appropriate to say that coherence resides at the textual, not the sentence level, and that it creates links between ideas to produce meaning and reveal organisation (Knoch, 2009). Moreover, coherence is not solely determined by the writer; the reader plays a part, as coherence is also concerned with how people receive and interpret a particular text (Knoch, 2009).

One of the most problematic aspects in assessing writing is coherence, due to its subjective nature (Todd, Thienpermpool & Keyuravong, 2004). Topical structural analysis and topic-based analysis are the most common measures used to examine coherence. Following is a brief discussion of topical structural analysis (TSA) and then topic-based analysis.

3.3.2.2.1 Topical structural analysis

TSA was developed by Lautamatti (1987). It analyses a sentence in two parts—a topic and a comment (Todd et al., 2004). The topic is the sentence’s main idea and the comment is what is said about the topic. Three types of topical progression are suggested by Lautamatti; these are outlined below (Todd et al., 2004):

1. Parallel progression (PP): the topic of the following sentence is the same as the preceding one. As this produces repetition of the topic, it reinforces an idea with the reader (<a b>, <a c>, <a, d>).
2. Sequential progression (SP): the topic of the successive sentence is a comment on the previous sentence or a different topic (<a b>, <b c>, <c d>).

3. Extended parallel progression (EPP): the topic is the same as in the previous sentence, but is interrupted with some SP (<a b>, <b c>, <a d>).

Regarding the effectiveness of TSA, past research has produced mixed results. Witte (1983) assessed high-rated and low-rated persuasive written compositions using TSA. Witte determined that high-rated persuasive written texts contained more PP and EPP and fewer SP. In contrast, Schneider and Connor (1990) analysed the progression of sentence topics in the *Test of Written English* (TWE); here, the results indicated that SP was used with higher level written composition, and that PP and EPP were used in the low- and medium-level essays. As for EPP, no differences were apparent in the three different levels. In a more recent study, Almaden (2006) examined topical progression using TSA in paragraphs written by Filipino ESL students. This study determined that PP was used most frequently, followed by EPP and SP. After using the original TSA in their study, Schneider and Connor (1990) proposed subcategories for SP:

1. Direct SP includes: a) neighbouring topics related by topic-commented patterns, b) word derivations (science, scientists), and c) part-whole relations (these groups, housewives, children and old people).

2. Indirectly related SP is related by semantic set (scientists, their inventions and discoveries, and the invention of the radio, telephone and television).

3. Unrelated SP shows no clear relation to either the previous sentence topic or the discourse topic (p. 422).

Knoch (2007, 2009) found that higher level essays used direct SP, indirect progression and superstructure more often. Conversely, lower level essays employed more PP, coherence breaks and unrelated progression. Both groups used extended progression equally.
3.3.2.2.2 *Topic-based analysis*

Topic-based analysis was originally developed by Watson Todd (1998) to analyse spoken classroom discourse. However, Watson Todd, Thienpermpool and Keyuravong (2004) used this measure to analyse written texts, as it allowed researchers to examine coherence more objectively. As they state:

Key concepts are identified primarily through frequency, the logical relationships between these concepts are identified, and from these relationships line diagrams representing schemata of the discourse are drawn up. The coherence of the discourse can then be measured by mapping the discourse onto these line diagrams. (p. 88)

Topic-based analysis involves six stages:

1. Stage 1—preparing for the text analysis
2. Stage 2—identifying key concepts
3. Stage 3—identifying relationships
4. Stage 4—linking the relationship into a hierarchy
5. Stage 5—mapping the discourse onto the hierarchy
6. Stage 6—identifying topics and measuring coherence.

Watson Todd et al. (2004) concluded that this method held promise for examining L2 learners’ written composition. However, Knoch, Macqueen et al. (2014) suggested that the process was complicated and laborious and could create difficulties when operationalising for statistical analysis.

As indicated by Watson Todd et al. (2004), coherence is subjective and problematic when assessing writing. The above assessment of past studies indicates mixed results with this measure. In addition, Knoch, Macqueen et al. (2014) suggested it was difficult to operationalise for statistical analysis. As such, coherence was not considered in the pilot study.
3.3.3 Desu/masu form and plain form

*Desu/masu* is a Japanese linguistic feature; and is often referred to as the ‘addressee honorific’. Many L2 learners prefer learning to speak before learning to write (Sato, 1987). This may cause problems when a learner begins using the *desu/masu* and plain forms in writing. To determine if overseas students studying in Japan were using the *desu/masu* and plain forms appropriately in their writing, Sato (1987) examined the written compositions of two students. One was Taiwanese, and the other was Korean. Sato discovered that both students wrote inconsistently; they used a mixture of the *desu/masu* and plain forms in written composition, which indicated that their sentences were not coherent. Sato (1987) suggested that learners who put more effort into speaking might produce mixed forms in their written compositions.

However, Sakai (2005) determined that uniformity in language was not as important as other factors. Sakai (2005) analysed the level of importance (as perceived by educators teaching in subjects other than Japanese) of different categories in a holistic rating scale. Fifty-one essays written by overseas students for scholarship applications to a Japanese university were examined in this study. Four factors were considered in the holistic rating scale: accuracy, richness, composition and content.

Sakai concluded that, out of the four assessment factors, content and richness were considered more important than accuracy and composition (*desu/masu* was included in the composition category). Sakai (2005) suggested that uniformity in the use of language was not nearly as important when assessing ‘good writing’, compared to other factors such as content and richness. For that reason, the *desu/masu* form was not examined as a measure in my pilot study.

### 3.4 The Chosen Discourse Analytic Measures

Based on the previous literature discussed in this section, a range of discourse analytic measures was chosen to be trialled in the pilot study (see Table 3.9).
Table 3.9

*Summary of Measures*

<table>
<thead>
<tr>
<th>Kanji complexity</th>
<th>1.</th>
<th>KC TTR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.</td>
<td>KC RTTR</td>
</tr>
<tr>
<td></td>
<td>3.</td>
<td>KC CTTR</td>
</tr>
<tr>
<td></td>
<td>4.</td>
<td>KW TTR</td>
</tr>
<tr>
<td></td>
<td>5.</td>
<td>KW RTTR</td>
</tr>
<tr>
<td></td>
<td>6.</td>
<td>KW CTTR</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Syntactic complexity</th>
<th>1.</th>
<th>Number of words per T-unit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.</td>
<td>Number of clauses per T-unit</td>
</tr>
<tr>
<td></td>
<td>3.</td>
<td>Words per sentence</td>
</tr>
<tr>
<td></td>
<td>4.</td>
<td>Clauses per sentence</td>
</tr>
<tr>
<td></td>
<td>5.</td>
<td>Ratio of te-form to words</td>
</tr>
<tr>
<td></td>
<td>6.</td>
<td>Ratio of simple sentences to sentences</td>
</tr>
<tr>
<td></td>
<td>7.</td>
<td>Ratio of complex sentences to sentences</td>
</tr>
<tr>
<td></td>
<td>8.</td>
<td>Ratio of clause conjunction to words</td>
</tr>
<tr>
<td></td>
<td>9.</td>
<td>Ratio of sentencing-ending forms to words</td>
</tr>
</tbody>
</table>

Sentence-ending forms include propositional endings, negative endings, past endings, explanatory endings, judgement endings, epistemic endings, expressive endings, nominal endings

<table>
<thead>
<tr>
<th>Accuracy</th>
<th>1.</th>
<th>Words per error-free sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.</td>
<td>Ratio of error-free sentence</td>
</tr>
<tr>
<td></td>
<td>3.</td>
<td>Ratio of error-free KCs</td>
</tr>
<tr>
<td></td>
<td>4.</td>
<td>Ratio of error-free KWs</td>
</tr>
<tr>
<td></td>
<td>5.</td>
<td>Ratio of error-free <em>kana</em></td>
</tr>
<tr>
<td></td>
<td>6.</td>
<td>Ratio of error-free particles</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fluency</th>
<th>1.</th>
<th>Number of words</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.</td>
<td>Number of clauses</td>
</tr>
<tr>
<td></td>
<td>3.</td>
<td>Number of sentences</td>
</tr>
<tr>
<td></td>
<td>4.</td>
<td>Number of KC</td>
</tr>
<tr>
<td></td>
<td>5.</td>
<td>Number of KW</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cohesion</th>
<th>1.</th>
<th>Ratio of general conjunctions to words</th>
</tr>
</thead>
</table>
Chapter 4: Study 1 Methodology

The research for this project comprised two studies: a cross-sectional study (Study 1) and a longitudinal classroom-based study (Study 2). Study 1’s objective was to determine how the writing ability of L2 Japanese learners at university varies across different levels. The current chapter presents a detailed description of Study 1’s methods. The chapter gives an overview of the study’s design, followed by a detailed description of the participants, instruments and procedures.

4.1 Study 1 Design

Study 1 was a cross-sectional study that investigated five different levels of writing scripts from students who participated in a Japanese placement test at a large metropolitan Australian university. At the beginning of Study 1, a pilot study was conducted to select the most suitable discourse analytic measures. The main aim of the pilot study was to identify which discourse analytic measures would distinguish writing at different levels of L2 Japanese proficiency. A larger number of written scripts were examined in the study’s main analysis phase, using the chosen discourse measures decided upon in the pilot study. The data analysis was quantitative and used statistical procedures.

Data for Study 1 were based on writing scripts collected in a trial of a placement test undertaken in Semester 1, 2013. Placement tests are conducted by the Japanese Department at the university every year in Semesters 1 and 2. The majority of students take Japanese as an elective subject, which means their Japanese study background varies. Some students have studied Japanese in secondary school, some have no experience and others are self-taught. This variation is the main reason for placement tests; in this way, students can be allocated to the appropriate class level before enrolling in a Japanese course.
One-hundred-and-forty-six writing scripts from a trial placement test were available for research purposes. Based on the placement test results, students were divided into five different levels: 1, 3, 5, 7 and 9. Level 1 signifies the lowest level and Level 9 the highest. The levels are all odd numbers (1, 3, 5) because in Semester 1, only odd numbered classes are offered; in Semester 2, only even numbered classes are offered. Therefore, those enrolled in Level 1, Semester 1 will be in Level 2 for Semester 2. As the placement test was conducted in Semester 1, only the written scripts from odd numbered classes were available.

The distribution of test papers from the different levels is shown in Table 4.1.

Table 4.1

<table>
<thead>
<tr>
<th>Level</th>
<th>Score</th>
<th>Student number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1</td>
<td>0–25</td>
<td>9</td>
</tr>
<tr>
<td>Level 3</td>
<td>26–77</td>
<td>40</td>
</tr>
<tr>
<td>Level 5</td>
<td>78–105</td>
<td>45</td>
</tr>
<tr>
<td>Level 7</td>
<td>106–130</td>
<td>34</td>
</tr>
<tr>
<td>Level 9</td>
<td>131–170</td>
<td>18</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>146</td>
</tr>
</tbody>
</table>

4.2 Participants

A background questionnaire completed by participants prior to the placement test revealed that students had either an Asian (Chinese being dominant) language or English as their first language. Only one student was from a Japanese background. A summary of the participants’ first languages is given in Table 4.2.

Table 4.2

<table>
<thead>
<tr>
<th>First language</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>80</td>
</tr>
<tr>
<td>Chinese (Mandarin)</td>
<td>37</td>
</tr>
<tr>
<td>Korean</td>
<td>6</td>
</tr>
<tr>
<td>First language</td>
<td>No.</td>
</tr>
<tr>
<td>----------------</td>
<td>-----</td>
</tr>
<tr>
<td>Cantonese</td>
<td>5</td>
</tr>
<tr>
<td>Vietnamese</td>
<td>5</td>
</tr>
<tr>
<td>Japanese</td>
<td>1</td>
</tr>
<tr>
<td>Other</td>
<td>12</td>
</tr>
<tr>
<td>(Slovak, Lao, Malay, Urdu, Filipino, Thai, Serbian, Shanghainese, Indonesian)</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>146</strong></td>
</tr>
</tbody>
</table>

The sample included more females \((n = 95)\) than males \((n = 51)\). The information on students’ ages was unavailable, but as most students were enrolled in the university’s undergraduate program, it is likely that most participants were in their late teens or early twenties. The majority of participants were from the Faculty of Arts \((n = 61)\) and the Faculty of Science \((n = 33)\). Others came from the Faculties of Business \((n = 22)\), Medicine \((n = 20)\), Architecture \((n = 7)\), Music \((n = 2)\) and Land and Environment \((n = 1)\).

### 4.3 Instruments

Two instruments were used: a written task and a background questionnaire.

#### 4.3.1 Placement test/writing task

The aim of writing task for the placement test was to examine how well participants could write in Japanese. They were asked to write an email in Japanese to a Japanese-speaking friend living in Japan and outline plans for a visit to Japan (see Appendix A). Participants were to write as much as possible within the time limit of 30 minutes.

#### 4.3.2 Background questionnaire

The background questionnaire was completed at the same time as the writing task. The questionnaire was divided into three sections (see Appendix B):

1. **Personal details**: name, country of birth, student ID, gender, residency status, course type, enrolment status, faculty, degree.
2. **Language background**: languages spoken, participant’s first language, questions relating to Japanese language (e.g., whether or not any member of their family spoke Japanese).

3. **Education**: questions relating to Japanese education background (e.g., whether or not the participant had studied Japanese in Japan).

### 4.4 Procedures

The written scripts were collected prior to the start of the study. In order to obtain access to the written scripts, ethical approval from the university was necessary. Once approval was obtained (see Appendix C), all 146 written and scored scripts were photocopied. The copies of the written scripts were coded and recorded in a database for easy reference.

#### 4.4.1 Pilot study

A pilot study was conducted prior to Study 1 to determine the discourse analytic measures used in the main study (Study 1). Based on the literature discussed in Chapter 3, a range of discourse analytic measures (complexity, accuracy, fluency and cohesion) were trialled in the pilot study. Five written scripts from each level (1, 3, 5, 7 and 9) with roughly mid-range scores were selected. The researcher conducted manual analysis of the pilot study. Due to the pilot study’s small sample size, no inferential statistics were obtained and data were not double-coded. Discourse analytic measures with a positive linear relationship (except for a few measures that were expected to demonstrate a negative linear relationship) between the measure and the proficiency levels were considered in Study 1. The means of each proficiency level were observed as able to determine the linear relationship between the measure and the proficiency level. The section below outlines the methods and pilot study results. It explains the data-coding procedure and the selection and rejection of certain measures for the main
study. The examples used in the data-coding procedure were taken from the written scripts used in the pilot study.

4.4.1.1 Data-coding procedure and pilot study results

4.4.1.1.1 Kanji complexity

Several kanji complexity measures were trialled in the pilot study. Table 4.3 illustrates the measures and the coding procedure with examples for KC TTR, KC RTTR, KC CTTR, KW TTR, KW RTTR and KW CTTR.

Table 4.3

<table>
<thead>
<tr>
<th>Measures</th>
<th>Coding</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>KC TTR</td>
<td>1) Count all the KC (token) a; 2) Count the KC excluding repeated ones (type) b; 3) Divide b by a (TTR).</td>
<td>私は日本に行った時に、東京に行きましたが、一日しか東京にいませんでした。</td>
</tr>
<tr>
<td>KC RTTR</td>
<td>( \sqrt{\frac{KCT}{KCN}} )</td>
<td>The TTR example above, there were 9 KC types and 12 KC tokens.</td>
</tr>
<tr>
<td>KC CTTR</td>
<td>( \sqrt{\frac{KCT}{2KCN}} )</td>
<td>As per above, there were 9 KC types and 12 KC tokens.</td>
</tr>
<tr>
<td>Measures</td>
<td>Coding</td>
<td>Examples</td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
<td>----------</td>
</tr>
<tr>
<td>3) Divide b by a (TTR).</td>
<td>When I went to Japan, I went to Tokyo but only stayed one day in Tokyo. Green + yellow = KW token (9) Green only = KW type = (7) TTR= ( \frac{7}{9} = 0.78 )</td>
<td></td>
</tr>
</tbody>
</table>

The TTR example above, there were 9 KW types and 7 KW tokens. RTTR = \( \frac{7}{\sqrt{9}} = \frac{7}{3} = 2.33 \)

As per above, there were 9 KW types and 7 KW tokens. CTTR = \( \frac{7}{\sqrt[3]{13}} = \frac{7}{\sqrt[3]{13}} = 1.65 \)

The results for kanji complexity measures are shown in Table 4.4, which displays the mean and standard deviations for each measure. The results are arranged by five different levels (1, 3, 5, 7 and 9). The measures considered for Study 1 were KC RTTR, KC CTTR, KW RTTR and KW CTTR. These all demonstrated a positive linear relationship between each measure and the level; in other words, the mean score increased as the level increased. KC TTR and KW TTR both showed a positive linear relationship from Level 1 to 7; however, they both declined at Level 9. As Park (2013) has noted, in the literature, RTTR and CTTR are reliable measures for complexity in comparison to TTR. As such, KC TTR and KW TTR were not considered in the main study.
Table 4.4

Results for Various Kanji Complexity Measures

<table>
<thead>
<tr>
<th>Measures</th>
<th>Level 1 (n = 5)</th>
<th>Level 3 (n = 5)</th>
<th>Level 5 (n = 5)</th>
<th>Level 7 (n = 5)</th>
<th>Level 9 (n = 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>KC TTR</td>
<td>0.50</td>
<td>0.48</td>
<td>0.55</td>
<td>0.10</td>
<td>0.64</td>
</tr>
<tr>
<td>KC RTTR</td>
<td>2.10</td>
<td>2.20</td>
<td>4.00</td>
<td>0.91</td>
<td>4.37</td>
</tr>
<tr>
<td>KC CTTR</td>
<td>1.48</td>
<td>1.56</td>
<td>2.83</td>
<td>0.64</td>
<td>3.09</td>
</tr>
<tr>
<td>KW TTR</td>
<td>0.51</td>
<td>0.48</td>
<td>0.59</td>
<td>0.08</td>
<td>0.67</td>
</tr>
<tr>
<td>KW RTTR</td>
<td>1.52</td>
<td>1.54</td>
<td>3.33</td>
<td>0.36</td>
<td>3.87</td>
</tr>
<tr>
<td>KW CTTR</td>
<td>1.08</td>
<td>1.10</td>
<td>2.33</td>
<td>0.27</td>
<td>2.59</td>
</tr>
</tbody>
</table>

4.4.1.1.2 Syntactic complexity

This section will examine syntactic complexity as two elements: syntactic measures and sentence-ending variables. Sentence-ending variables are classified under syntactic complexity. However, as a number of measures for sentence-ending variables were trialled in the pilot study, these measures will be reported after the measures for syntactic complexity. The measures deemed appropriate for syntactic complexity after reviewing the literature were trialled in the pilot study. The coding procedure and examples for syntactic measures are shown in Table 4.5.

Table 4.5

Syntactic Complexity Measures

<table>
<thead>
<tr>
<th>Measures</th>
<th>Coding</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of words</td>
<td>A T-unit is a main clause with all the subordinate clauses (Hunt, 1970). T-unit boundaries were coded following the coding scheme set out in Cumming et al. (2006) and were adapted for this study (see Appendix D)</td>
<td>日本には一ヶ月ほどいるつもりですが、多分東京の方に二週間ぐらいいると思います。I intend to stay in Japan for about a month but I think I will be in Tokyo for about 2 weeks. There is 1 T-unit in this</td>
</tr>
<tr>
<td>per T-unit</td>
<td>Total number of words Total number of Tunits</td>
<td></td>
</tr>
<tr>
<td>Measures</td>
<td>Coding</td>
<td>Examples</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Number of clauses Per T-unit</td>
<td>A clause is a string of words that contains a subject and a finite verb (Cummings et al., 2006)</td>
<td>I intend to stay in Japan for about a month but I think I will be in Tokyo for about 2 weeks. There are two clauses in this T-unit</td>
</tr>
<tr>
<td></td>
<td>T-units and clause boundaries were coded following the coding scheme set out in Cumming et al. (2006). See Appendix D</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Total number of clauses</strong> / <strong>Total number of T-units</strong></td>
<td></td>
</tr>
<tr>
<td>Words per sentence</td>
<td>As Japanese words encompasses content and grammatical words both were counted. Numbers were counted as one word.</td>
<td>Is Japan hot?</td>
</tr>
<tr>
<td></td>
<td><strong>Total number of Japanese words in this sentence</strong></td>
<td>There are five Japanese words in this sentence</td>
</tr>
<tr>
<td>Clauses per sentence</td>
<td>Number of clauses per sentence was counted. For clause boundaries, see Appendix D</td>
<td>I intend to stay in Japan for about a month but I think I will be in Tokyo for about 2 weeks. There are two clauses in this sentence.</td>
</tr>
<tr>
<td></td>
<td><strong>Total number of clauses in this sentence</strong></td>
<td></td>
</tr>
<tr>
<td>Ratio of te-form to words</td>
<td>Various te-form were counted and divided by total number of words to obtain the ratio.</td>
<td>I will look into various tourist destinations and will make my plan.</td>
</tr>
<tr>
<td></td>
<td><strong>Total number of te-form</strong> / <strong>Total number of words</strong></td>
<td></td>
</tr>
<tr>
<td>Ratio of simple sentences</td>
<td>Simple sentences contain one verb and one main idea (Iwashita &amp; Sekiguchi, 2009). Simple sentences were counted and divided by total number of sentences.</td>
<td>I will go to Japan next month.</td>
</tr>
<tr>
<td></td>
<td><strong>Total number of simple sentences</strong> / <strong>Total number of sentences</strong></td>
<td></td>
</tr>
</tbody>
</table>
**Measures** | **Coding** | **Examples**
--- | --- | ---
Ratio of complex sentences | Complex sentences contain two or more clauses, two or more verbs and more than one main idea. Complex sentences were counted and divided by total number of sentences.

\[
\text{Total number of complex sentences} \quad \frac{\text{Total number of sentences}}{}
\]

Although I have heard about this before but is it wrong to wear shoes inside the house?

Ratio of clause conjunctions to words | Clause conjunction was counted and divided by total number of words to obtain the ratio.*

\[
\text{Total number of clause conjunctions} \quad \frac{\text{Total number of words}}{}
\]

私は新幹線に乗って、ゆきさんと一緒に東京へ行きたいです。

Watashi wa shinkansen ni note Yuki san to issho ni Tokyo e ikitaidesu.

*Note: Tanaka (1997) classified clause conjunction into six different categories; conjunctive particles, conjugal forms, relativisation forms, nominalisers, quotation particles and nominal endings. In this study, these were all classified as ‘clause conjunctions’ and were not divided into each category.

The results for various syntactic complexity measures are shown in Table 4.6. Out of the eight measures trialled for syntactic complexity, the results suggest that only the ratio of te-form to words, the ratio of complex sentences and the ratio of clause conjunctions were successful measures, as these showed a positive linear relationship between the measure and the levels. The ratio of te-form per words did not increase from Level 5 to 7 but showed an increase from Level 7 to 9; hence, it was deemed appropriate to include in Study 1. The remaining five measures showed an increase from Level 1 to 3, but most showed a negative linear relationship from Level 3 onwards and therefore were not included in Study 1.
Table 4.6

Results for the Syntactic Complexity Measures

<table>
<thead>
<tr>
<th>Measures</th>
<th>Level 1 (n = 5)</th>
<th>Level 2 (n = 5)</th>
<th>Level 3 (n = 5)</th>
<th>Level 4 (n = 5)</th>
<th>Level 5 (n = 5)</th>
<th>Level 6 (n = 5)</th>
<th>Level 7 (n = 5)</th>
<th>Level 8 (n = 5)</th>
<th>Level 9 (n = 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>No. of words per T-unit</td>
<td>0.00</td>
<td>0.00</td>
<td>12.20</td>
<td>4.70</td>
<td>12.00</td>
<td>3.00</td>
<td>13.40</td>
<td>1.14</td>
<td>12.00</td>
</tr>
<tr>
<td>No. of clauses per T-unit</td>
<td>0.00</td>
<td>0.00</td>
<td>2.00</td>
<td>0.71</td>
<td>2.00</td>
<td>0.00</td>
<td>2.20</td>
<td>0.45</td>
<td>2.00</td>
</tr>
<tr>
<td>Words per sentence</td>
<td>4.90</td>
<td>1.83</td>
<td>8.55</td>
<td>1.42</td>
<td>8.26</td>
<td>1.63</td>
<td>11.95</td>
<td>2.70</td>
<td>9.08</td>
</tr>
<tr>
<td>Clauses per sentence</td>
<td>1.00</td>
<td>1.00</td>
<td>1.26</td>
<td>0.09</td>
<td>1.14</td>
<td>0.05</td>
<td>1.75</td>
<td>0.51</td>
<td>1.41</td>
</tr>
<tr>
<td>Ratio of te-form per words</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>1.82</td>
<td>0.02</td>
<td>1.67</td>
<td>0.02</td>
<td>1.92</td>
<td>0.36</td>
</tr>
<tr>
<td>Ratio of simple sentence to sentences</td>
<td>0.34</td>
<td>1.64</td>
<td>0.55</td>
<td>2.88</td>
<td>0.71</td>
<td>2.77</td>
<td>0.41</td>
<td>4.60</td>
<td>0.61</td>
</tr>
<tr>
<td>Ratio of complex sentence to sentences</td>
<td>0.00</td>
<td>0.00</td>
<td>0.11</td>
<td>1.48</td>
<td>0.13</td>
<td>1.00</td>
<td>0.34</td>
<td>2.83</td>
<td>0.36</td>
</tr>
<tr>
<td>Ratio of clause conjunction to words</td>
<td>0.00</td>
<td>0.00</td>
<td>4.80</td>
<td>3.35</td>
<td>5.00</td>
<td>3.39</td>
<td>10.40</td>
<td>4.83</td>
<td>11.20</td>
</tr>
</tbody>
</table>

4.4.1.1.3 Syntactic complexity—Ratio of sentence-ending variables to words

As mentioned in Section 3.2.1.3.4, sentence-ending variables are considered important when examining the naturalness of writing in Japanese (Tanaka, 1997). They were categorised under syntactic complexity. Sentence-ending variables adopted from Tanaka (1997) were trialled in the pilot study. Table 4.7 shows the sentence-ending variables and examples (brackets are English translations). The coding procedure and an example are shown after Table 4.7.
Table 4.7  

Sentence-ending Variables and Examples

<table>
<thead>
<tr>
<th>Sentence-ending variables</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Propositional endings</td>
<td><em>suru</em> (do), <em>da</em> (is);</td>
</tr>
<tr>
<td>Negative endings</td>
<td><em>shinai</em> (do not), <em>dewanai</em> (is not)</td>
</tr>
<tr>
<td>Past endings</td>
<td><em>shita</em> (did), <em>datta</em> (was)</td>
</tr>
<tr>
<td>Explanatory endings</td>
<td><em>suru no da</em> (do), <em>suru wake da</em> (so it means)</td>
</tr>
<tr>
<td>Judgement endings</td>
<td><em>surubeki da</em> (should do)</td>
</tr>
<tr>
<td></td>
<td><em>shitahogayoi</em> (had better do)</td>
</tr>
<tr>
<td>Epistemic endings</td>
<td><em>to omou</em> (think that)</td>
</tr>
<tr>
<td></td>
<td><em>surukamoshirenai</em> (might do)</td>
</tr>
<tr>
<td>Expressive endings</td>
<td><em>shitai</em> (want to do), <em>shiyo</em> (let’s do)</td>
</tr>
<tr>
<td>Nominal endings</td>
<td><em>da/dearu</em> (is)</td>
</tr>
</tbody>
</table>

The occurrence of each sentence-ending variable was counted for each written script. To obtain the ratio, the number of sentence-ending variables was then divided by the total number of words in a script. The sentence-ending variables in the example given below are underlined.

Example:

日本に行きたい。なぜならこの一年の間でアニメに夢中

*ikitai*. *Nazenara konoichinenkan no aida de anime ni mucchu ni natta.*

I want to go to Japan. Because during this year I became so into animation.

Note: The example sentence is taken directly from the written script. Therefore, the sentence is not necessarily correct and is translated as expressed.

In this example, the sentence ending in the first sentence is an expressive ending, ‘*ikitai*’ (‘want to go’) and in the second sentence it is a past ending, ‘*natta*’ (‘became’). Sentence-ending variables were individually examined in terms of counting, which means this example contains one expressive ending and one past ending.

Table 4.8 shows the results for various sentence-ending variables. Epistemic endings, expressive endings and propositional endings were considered appropriate for
Study 1. Both epistemic endings and expressive endings showed no constant increase from level to level (e.g., expressive endings stayed constant from Level 3 to 7). Despite this, they were considered in Study 1 as both measures did show a gradual increase from Level 1 to 9. The ratio of propositional endings showed a negative linear relationship between the measure and the levels—with this particular measure, this trend should occur if a learner’s proficiency developed, as indicated in Tanaka’s (1997) study. A negative relationship was anticipated for this measure; therefore, it was retained for Study 1. Other measures stayed constant throughout or showed a negative linear relationship and hence were disregarded.

Table 4.8

Results for Various Sentence-ending Variables

<table>
<thead>
<tr>
<th>Measures (all in ratios per words)</th>
<th>Level 1 (n = 5)</th>
<th>Level 3 (n = 5)</th>
<th>Level 5 (n = 5)</th>
<th>Level 7 (n = 5)</th>
<th>Level 9 (n = 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Propositional endings</td>
<td>0.15</td>
<td>1.82</td>
<td>0.08</td>
<td>3.29</td>
<td>0.07</td>
</tr>
<tr>
<td>Negative endings</td>
<td>0.01</td>
<td>0.41</td>
<td>0.01</td>
<td>0.55</td>
<td>0.01</td>
</tr>
<tr>
<td>Past endings</td>
<td>0.01</td>
<td>0.45</td>
<td>0.01</td>
<td>0.89</td>
<td>0.01</td>
</tr>
<tr>
<td>Explanatory endings</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>Judgement endings</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Epistemic endings</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>1.79</td>
<td>0.02</td>
</tr>
<tr>
<td>Expressive endings</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>1.67</td>
<td>0.01</td>
</tr>
<tr>
<td>Nominal endings</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.45</td>
<td>0.00</td>
</tr>
</tbody>
</table>

4.4.1.4 Accuracy measures

After reviewing the literature, a number of accuracy measures were trialled in the pilot study. These measures are shown in Table 4.9, which also illustrates the coding procedure with an example for each measure.
### Table 4.9

**Accuracy Measures**

<table>
<thead>
<tr>
<th>Measures</th>
<th>Coding</th>
<th>Examples</th>
</tr>
</thead>
</table>
| Ratio of error-free sentences   | An error-free sentence is one for which there are no errors (KC, KW, particles, grammar etc). | 今/日本/の/天気/は/いい/ですか/か。
Is the weather in Japan good at the moment? The above example is an error-free sentence |
|                                 | Total no of error free sentences                                      | Total no of sentences                                                   |
|                                 | Total no of sentences                                                 |                                                                         |
| Ratio of EFCs                   | An error-free clause is one for which there are no errors (KC, KW, particles, grammar etc). | ところで、私は日本に行くつもりです。
By the way, I plan to go to Japan. The above example is an error-free clause |
|                                 | Total no of error free clauses                                        | Total no of clauses                                                     |
|                                 | Total no of clauses                                                   |                                                                         |
| Words per error-free sentence   | To count all the words in an error-free sentence, both content words and grammatical words were counted | 今/日本/の/天気/は/いい/ですか/か。
Is the weather in Japan good at the moment? |
|                                 | Total no of error free KC                                             | Total no of KC                                                          |
|                                 | Total no of KC                                                        |                                                                         |
| Ratio of error-free KC          | Coding of kanji errors adapted from Kubota (2005) See Appendix E.     | Next month on a Monday morning, I will be getting on a train at the Narita airport |
|                                 | Total no of error free KC                                             | Green + yellow = Total KC (14)
Green = KC error (3)
Error-free KC = 14 – 3 = 11
Ratio of error-free KC = \( \frac{11}{14} \) = 0.79 |

127
<table>
<thead>
<tr>
<th>Measures</th>
<th>Coding</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratio of error-free KW</td>
<td>$\frac{\text{Total no of error free KW}}{\text{Total no of KW}}$</td>
<td>Green + yellow = Total KW (7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Green = KW error (2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Error-free KW = $7 - 2 = 5$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ratio of error-free KW = $\frac{5}{7} = 0.71$</td>
</tr>
<tr>
<td>Coding of kanji errors adapted</td>
<td></td>
<td></td>
</tr>
<tr>
<td>See Appendix E.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>from Kubota (2005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>I want to go to Harajyuku and YoyogiKoen.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ratio of error-free Kana (KA)</td>
<td>$\frac{\text{Total no of error free KA}}{\text{Total no of KA}}$</td>
<td>Green + yellow = Total KA (20)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Green = KA error (2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Error-free KA = $7 - 2 = 5$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ratio of error-free KA = $\frac{18}{20} = 0.9$</td>
</tr>
<tr>
<td>Coding of kana errors adapted</td>
<td></td>
<td></td>
</tr>
<tr>
<td>See Appendix E.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>from Kubota (2005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>I want to go to a shop that sells a lot of cute stuff. Then, I want to</td>
</tr>
<tr>
<td></td>
<td></td>
<td>eat a delicious Japanese dish.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ratio of error-free particles</td>
<td>$\frac{\text{Total no of error free KA}}{\text{Total no of KA}}$</td>
<td>Green + yellow = Total P (4)</td>
</tr>
<tr>
<td>(P)</td>
<td></td>
<td>Green = P error (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Error-free P = $4 - 1 = 3$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ratio of error-free P = $\frac{3}{4} = 0.75$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As the results in Table 4.10 reveal, not all measures showed a substantial change between levels. The results show clearly that the ratio of error-free sentences and the ratio of EFCs had a positive linear relationship. Words per error-free sentence dropped slightly from Level 1 to 3, but the decrease was minimal. From Level 3 and above, a
relatively strong increase was apparent; as such, this measure was retained for Study 1. The results indicate that particle accuracy, KC accuracy, KW accuracy and KA accuracy all had high ratios of accuracy across all levels, which indicates that students only used the features of Japanese of which they had knowledge. These measures were not retained for Study 1, as they did not show any positive linear relationship.

Table 4.10

Results for Various Accuracy Measures

<table>
<thead>
<tr>
<th>Measures</th>
<th>Level 1 (n = 5)</th>
<th>Level 3 (n = 5)</th>
<th>Level 5 (n = 5)</th>
<th>Level 7 (n = 5)</th>
<th>Level 9 (n = 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Ratio of error-free sentences</td>
<td>0.21</td>
<td>1.52</td>
<td>0.22</td>
<td>2.12</td>
<td>0.25</td>
</tr>
<tr>
<td>Ratio of EFCs</td>
<td>0.12</td>
<td>2.23</td>
<td>0.23</td>
<td>2.54</td>
<td>0.27</td>
</tr>
<tr>
<td>Words per error-free sentence</td>
<td>5.94</td>
<td>2.87</td>
<td>5.30</td>
<td>3.27</td>
<td>7.12</td>
</tr>
<tr>
<td>Ratio of error-free KC</td>
<td>0.98</td>
<td>2.49</td>
<td>0.98</td>
<td>2.05</td>
<td>0.96</td>
</tr>
<tr>
<td>Ratio of error-free KW</td>
<td>0.97</td>
<td>4.67</td>
<td>0.96</td>
<td>2.20</td>
<td>0.91</td>
</tr>
<tr>
<td>Ratio of error-free KA</td>
<td>0.96</td>
<td>4.18</td>
<td>0.97</td>
<td>1.30</td>
<td>0.91</td>
</tr>
<tr>
<td>Ratio of error-free particles</td>
<td>0.93</td>
<td>9.10</td>
<td>0.86</td>
<td>6.80</td>
<td>0.90</td>
</tr>
</tbody>
</table>

4.4.1.1.5 Fluency measures

Fluency measures trialled in the pilot study were the number of words (overall), the number of sentences, the number of clauses, the number of KCs and the number of KWs. Fluency measures and their coding procedures are given in Table 4.11. Examples for fluency are not shown in this table, as the procedures were simple (e.g., counting the number of words).
Table 4.11

*Fluency Measures*

<table>
<thead>
<tr>
<th>Measures</th>
<th>Coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of words</td>
<td>Coding of the number of words was undertaken by counting all the words (both content and grammatical words in the written script)</td>
</tr>
<tr>
<td>Number of sentences</td>
<td>Coding of the number of sentences was undertaken by counting all the sentences in the written script</td>
</tr>
<tr>
<td>Number of clauses</td>
<td>Coding of the number of clauses was undertaken by counting all the clauses in the written script</td>
</tr>
<tr>
<td>Number of KC</td>
<td>Coding of the KC was undertaken by counting the number of KC. The same procedure as explained for counting KC (token) in TTR was used</td>
</tr>
<tr>
<td>Number of KW</td>
<td>Coding of the KW was undertaken by counting the numbers of KW. The same procedure as explained for counting KW (token) in TTR was used</td>
</tr>
</tbody>
</table>

The results for the fluency analysis are given in Table 4.12. All the measures were included for Study 1, as they demonstrated a positive linear relationship between each measure and the level. In other words, the mean score increased as the level increased. For both the number of words and the number of clauses, a slight drop in the mean number from between Level 7 and 9 was evident. However, these two measures were retained. First, there was a strong increase in the mean number for the other levels and second, the small sample size could mean that the decrease in the highest level might have been coincidental.
Table 4.12

Results for Various Fluency Measures

<table>
<thead>
<tr>
<th>Measures</th>
<th>Level 1 $(n = 5)$</th>
<th>Level 3 $(n = 5)$</th>
<th>Level 5 $(n = 5)$</th>
<th>Level 7 $(n = 5)$</th>
<th>Level 9 $(n = 5)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
</tr>
<tr>
<td>Number of words</td>
<td>30.0</td>
<td>12.9</td>
<td>109.0</td>
<td>20.4</td>
<td>134.6</td>
</tr>
<tr>
<td>Number of sentences</td>
<td>6.4</td>
<td>2.88</td>
<td>13.4</td>
<td>2.07</td>
<td>17.2</td>
</tr>
<tr>
<td>Number of clauses</td>
<td>4.80</td>
<td>2.68</td>
<td>16.00</td>
<td>2.74</td>
<td>16.60</td>
</tr>
<tr>
<td>Number of KC</td>
<td>15.2</td>
<td>20.6</td>
<td>55.4</td>
<td>18.78</td>
<td>54.40</td>
</tr>
<tr>
<td>Number of KW</td>
<td>7.40</td>
<td>9.79</td>
<td>33.60</td>
<td>9.50</td>
<td>36.80</td>
</tr>
</tbody>
</table>

4.4.1.1.6 Cohesion measure: Ratio of general conjunctions to words

The cohesion measure trialled for the pilot study was the ratio of general conjunctions to words. The coding procedure and an example are shown below.

Occurrences of general conjunctions in the written scripts were counted following Tanaka’s (1997) coding procedure. To obtain the ratio, the number of general conjunctions was then divided by the total number of words in a script. General conjunctions are conjunctions such as *dakara* (‘therefore’), which relate two independent sentences without making a complex sentence.

Example:

前友達と東京であそびました。**だから** 次は家族と関西

行くつもりです

*Mae tomodachito Tokyo de asobimashita. **Dakara** tsugi wa kazoku to kansai e

ikutsumoridesu.*

Before I played with my friend in Tokyo. **Therefore** next time I plan to go to Kansai with my family.
Note: The example sentence is taken directly from the written script. Therefore, the sentence is not necessarily correct and is translated as expressed.

_Dakara_ (‘therefore’) was used to connect two sentences in this example. This was counted as one general conjunction in a written script.

The cohesion analysis results are given in Table 4.13. The results show that the mean value dropped significantly between Level 7 and 9 for the ratio of general conjunctions to words. Tanaka (1997) compared the development of general and clause conjunctions. He suggested that as the level increased, the number of general conjunction decreased, as the number of clause conjunction increased. Therefore, although a decrease in general conjunctions did occur from Level 7 to 9, this was still considered in Study 1.

Table 4.13

_Cohesion Results_

<table>
<thead>
<tr>
<th>Measures</th>
<th>Level 1 (n = 5)</th>
<th>Level 3 (n = 5)</th>
<th>Level 5 (n = 5)</th>
<th>Level 7 (n = 5)</th>
<th>Level 9 (n = 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Ratio of general conjunctions to words</td>
<td>0.60</td>
<td>1.34</td>
<td>2.40</td>
<td>0.89</td>
<td>4.80</td>
</tr>
</tbody>
</table>

Based on the pilot study reported above, the measures chosen for Study 1 are shown in Table 4.14.
Table 4.14

Measures Used in Study 1

<table>
<thead>
<tr>
<th>Measures Used in Study 1</th>
<th>Kanji complexity</th>
<th>Syntactic complexity</th>
<th>Accuracy</th>
<th>Fluency</th>
<th>Cohesion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1. KC RTTR</td>
<td>1. Ratio of te-form</td>
<td>1. Ratio of error-free sentences</td>
<td>1. Number of words</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. KC CTTR</td>
<td>2. Ratio of complex sentences to sentences</td>
<td>2. Ratio of EFCs</td>
<td>2. Number of clauses</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. KW RTTR</td>
<td>3. Ratio of clause conjunction to words</td>
<td>3. Words per error-free sentence</td>
<td>3. Number of sentences</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4. KW CTTR</td>
<td>4. Ratio of propositional endings to words</td>
<td></td>
<td>4. Number of KC</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. Ratio of epistemic endings to words</td>
<td></td>
<td>5. Number of KW</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6. Ratio of expressive endings to words</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.4.2 Statistical analysis

As only five written scripts for each level (25 written scripts in total) were used in the pilot study, no statistical tests were conducted. However, data analysis was conducted in Study 1, using the statistical package SPSS. The analysis was undertaken in several parts. First, descriptive statistics were obtained using SPSS for each of the measures considered in this study. For those measures that violated the assumption of constant variance, a log transformation was used. Second, analysis of variance (ANOVA) tests were performed. Finally, post-hoc least significant difference (LSD) tests were conducted to compare the significant differences between each level where such differences were found. Effect sizes were also considered when reporting the
results. Based on Cohen (1998), the $\eta^2$ (Eta squared) effect sizes were interpreted as follows: .01 = small; .06 = medium; .14 = large.

4.4.3 Inter-coder reliability analysis for Study 1

In the pilot study, the researcher coded measures and developed coding guidelines. Before the analysis of Study 1 was conducted, approximately ten per cent of the writing scripts (14 out of 146) were randomly selected and double-coded to ensure inter-coder reliability. A Japanese teacher who is a native Japanese speaker was chosen as a second coder. To ensure understanding of the concepts (e.g., kanji complexity, fluency and accuracy), the second coder was trained through coding two or three scripts, reviewing the coding guidelines and discussing the coding procedure with the researcher.

In order to measure such pairwise correlations, Spearman’s Rho was employed using SPSS. Spearman’s Rho measures the degree of consistency between two raters. As the data measured were ratio data, this approach was appropriate to estimate inter-rater reliability.

The results from Spearman’s Rho showed a high rate of inter-rater reliability. Tables 4.15, 4.16, 4.17, 4.18 and 4.19 show these correlation coefficients for each measure.

The Spearman’s Rho statistics for all measures were above .90. The percentages of agreement for the measures were all above 90%. This indicates that the measures were all reliable.

Table 4.15

*Kanji Complexity Inter-rater Reliability*

<table>
<thead>
<tr>
<th></th>
<th>KC RTTR</th>
<th>KC CTTR</th>
<th>KW RTTR</th>
<th>KW CTTR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman’s Rho</td>
<td>.91</td>
<td>.91</td>
<td>.97</td>
<td>.97</td>
</tr>
</tbody>
</table>
Table 4.16

*Syntactic Complexity Inter-rater Reliability*

<table>
<thead>
<tr>
<th></th>
<th>Te-form</th>
<th>Complex sentences</th>
<th>Clause conjunction</th>
<th>Propositional endings</th>
<th>Epistemic endings</th>
<th>Expressive endings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman’s Rho</td>
<td>1.00</td>
<td>.98</td>
<td>1.00</td>
<td>1.00</td>
<td>.91</td>
<td>.91</td>
</tr>
</tbody>
</table>

Table 4.17

*Accuracy Inter-rater Reliability*

<table>
<thead>
<tr>
<th></th>
<th>Error-free sentences</th>
<th>Error-free clauses</th>
<th>Words per error-free sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman’s Rho</td>
<td>.98</td>
<td>.97</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 4.18

*Fluency Inter-rater Reliability*

<table>
<thead>
<tr>
<th></th>
<th>No. of words</th>
<th>No. of sentences</th>
<th>No. of clauses</th>
<th>No. of KC</th>
<th>No. of KW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman’s Rho</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 4.19

*Cohesion Inter-rater Reliability*

<table>
<thead>
<tr>
<th></th>
<th>General conjunctions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman’s Rho</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Chapter 5: Study 1 Results and Discussion

In order to explore whether the writing ability of L2 Japanese learners at university varied across different levels (RQ1), various writing features (comprising a number of measures for each dimension) were investigated in Study 1. First, all measures chosen from the pilot study for Study 1 were correlated with the levels to determine if a statistically significant correlation existed between level and measure. All measures that had a statistically significant relationship with the proficiency levels were used for the Study 1 analysis. A one-way ANOVA test was performed for each measure to determine if a statistically significant difference existed between the measure and the proficiency levels. To determine where the exact differences were located between the proficiency levels, post-hoc LSD tests were also conducted for those measures with significant differences in the overall ANOVA test.

The Study 1 results are divided into two sections. First, a summary of the correlations between the measures and the proficiency level is discussed. Second, the results from the ANOVA test for each measure are presented. In accordance with RQ1, the extent to which each measure demonstrated developmental changes will be discussed.

5.1 Statistical Analysis: Correlations

In order to examine if a statistically significant relationship existed between each of the measures and the level, a bivariate Pearson correlation test between each of the measures and proficiency level was performed.

Table 5.1 shows the results of the correlations. Most measures showed a positive moderate correlation between each of the measures and the proficiency level. This indicates that as the level of proficiency increased, so did the measures. Measures that showed a weak positive correlation between the proficiency levels were the ratio of epistemic endings, the ratio of expressive endings and the ratio of general conjunctions
to words. The ratio of propositional endings showed a negative moderate correlation, which was a predictable exception (see Chapter 3, 3.2.1.3.4 for more detail). All the measures investigated showed statistical significance at alpha $p < .05$, except for the ratio of general conjunctions to words, which did not increase with the proficiency level. As the ratio of general conjunctions to words was not statistically significant, this measure was excluded from the Study 1 analysis. In the pilot study, this measure was retained, as the subsample size was small ($n = 5$) and a definite conclusion could not be made. However, in this analysis, the sample size was sufficient ($N = 146$) to conclude that it was not an appropriate measure for Study 1’s investigation.

Table 5.1

*Results of Bivariate Pearson Correlation Tests*

<table>
<thead>
<tr>
<th>Measures</th>
<th>Correlation with level</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Kanji complexity</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KC RTTR</td>
<td>.64</td>
<td>$p &lt; .001$</td>
</tr>
<tr>
<td>KC CTTR</td>
<td>.64</td>
<td>$p &lt; .001$</td>
</tr>
<tr>
<td>KW RTTR</td>
<td>.66</td>
<td>$p &lt; .001$</td>
</tr>
<tr>
<td>KW CTTR</td>
<td>.63</td>
<td>$p &lt; .001$</td>
</tr>
<tr>
<td><em>Syntactic complexity</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ratio of te-form to words</td>
<td>.53</td>
<td>$p &lt; .001$</td>
</tr>
<tr>
<td>Ratio of complex sentences to sentences</td>
<td>.52</td>
<td>$p &lt; .001$</td>
</tr>
<tr>
<td>Ratio of clause conjunctions to words</td>
<td>.62</td>
<td>$p &lt; .001$</td>
</tr>
<tr>
<td>Ratio of propositional endings to words</td>
<td>-.42</td>
<td></td>
</tr>
<tr>
<td>Ratio of epistemic endings to words</td>
<td>.25</td>
<td>$p = .002$</td>
</tr>
<tr>
<td>Ratio of expressive endings to words</td>
<td>.19</td>
<td>$p = .021$</td>
</tr>
<tr>
<td><em>Accuracy</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ratio of error-free sentences</td>
<td>.50</td>
<td>$p &lt; .001$</td>
</tr>
<tr>
<td>Ratio of EFCs</td>
<td>.53</td>
<td>$p &lt; .001$</td>
</tr>
<tr>
<td>Words per error-free sentence</td>
<td>.52</td>
<td>$p &lt; .001$</td>
</tr>
<tr>
<td>Measures</td>
<td>Correlation with level</td>
<td>Significance</td>
</tr>
<tr>
<td>--------------------------</td>
<td>------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Fluency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of words</td>
<td>.66</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td>Number of sentences</td>
<td>.49</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td>Number of clauses</td>
<td>.75</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td>Number of KCs</td>
<td>.57</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td>Number of KWs</td>
<td>.60</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td>Cohesion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ratio of general ...</td>
<td>.03</td>
<td>p = .720</td>
</tr>
</tbody>
</table>

**5.2 Results and Discussion**

Study 1 investigated several writing features: kanji complexity, syntactic complexity, accuracy and fluency. Each of these features comprised a number of measures (explained in more detail below). For each measure investigated, a table shows the means and standard deviations for each proficiency level. Then, side-by-side box plots reveal the distribution over the different writing proficiency levels. The box of each plot represents the middle 50 per cent of students and the black line inside the box represents the median. The whiskers represent the lowest and highest 25 per cent of cases. Cases that lie above or below the whiskers are outliers. Third, the results obtained from a one-way ANOVA test are reported. Measures that did not meet equality of variance assumptions were log transformed before performing the ANOVA test. In order to establish where the significant differences were located between the proficiency levels, a post-hoc LSD test was further conducted for those measures that showed significant differences between the levels. The results of the post-hoc test are also presented. Finally, at the end of each section, the outcome of each measure is discussed briefly.
5.2.1 **Kanji complexity**

To determine whether learners could use a wide range of KCs and KWs in a written script, four measures were chosen as part of *kanji* complexity: KC RTTR, KC CTTR, KW RTTR and KW CTTR. For all measures of *kanji* complexity, the assumption of equality of variances was not violated and an ANOVA test was conducted. To examine where the statistically significant differences were present between the levels, a post-hoc LSD was conducted.

### 5.2.1.1 **Kanji character root type token ratio**

Table 5.2 displays the descriptive statistics for mean scores and standard deviations of KC RTTR among students at each proficiency level. Box plots illustrating this relationship are shown in Figure 5.1. These indicate a large jump in the mean between Level 1 and 3, and a gradual increase between Level 3 and 9; however, as Figure 5.1 shows, the range of responses largely overlapped between Level 3 and 9.

**Table 5.2**

**Descriptive Statistics—Kanji Complexity: Kanji Character Root Type Token Ratio**

<table>
<thead>
<tr>
<th>Level</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.82</td>
<td>1.61</td>
<td>.00</td>
<td>4.62</td>
</tr>
<tr>
<td>(n = 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3.95</td>
<td>.95</td>
<td>2.01</td>
<td>5.52</td>
</tr>
<tr>
<td>(n = 40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4.48</td>
<td>1.02</td>
<td>2.40</td>
<td>6.64</td>
</tr>
<tr>
<td>(n = 45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>5.52</td>
<td>1.05</td>
<td>3.57</td>
<td>7.58</td>
</tr>
<tr>
<td>(n = 34)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>5.84</td>
<td>1.42</td>
<td>3.26</td>
<td>8.26</td>
</tr>
<tr>
<td>(n = 18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In order to establish if any statistically significant differences existed between proficiency levels, a one-way ANOVA test was performed. The ANOVA test showed a statistically significant result between proficiency levels with a large effect size:

$$F(4, 141) = 29.28, \ p<.001, \ \eta^2 = 0.45.$$  

A post-hoc LSD test was conducted to test the differences in KC RTTR between each level and the next. The test revealed that a significant difference in KC RTTR was present between all levels except between Levels 7 and 9.

The results suggest that on average, learners were able to use a wider variety of KCs in their writing as their proficiency developed. A statistically significant increase in the average number of KC RTTR was observed at each level of proficiency, except for the highest Levels 7 and 9.

5.2.1.2 Kanji character corrected type token ratio

The descriptive statistics in Table 5.3 and the box plots in Figure 5.2 show that the mean number for KC CTTR increased as the level of proficiency increased. As with KC RTTR, there is a substantial overlap in the range of KC CTTR once students reach Level 3 (see Figure 5.2).
Table 5.3  
Descriptive Statistics—Kanji Complexity: Kanji Character Corrected Type Token Ratio

<table>
<thead>
<tr>
<th>Level</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.29</td>
<td>1.14</td>
<td>.00</td>
<td>3.27</td>
</tr>
<tr>
<td>(n = 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2.80</td>
<td>0.67</td>
<td>1.42</td>
<td>3.90</td>
</tr>
<tr>
<td>(n = 40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3.15</td>
<td>0.70</td>
<td>1.69</td>
<td>4.69</td>
</tr>
<tr>
<td>(n = 45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>3.90</td>
<td>0.74</td>
<td>2.53</td>
<td>5.36</td>
</tr>
<tr>
<td>(n = 34)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>4.13</td>
<td>1.00</td>
<td>2.30</td>
<td>5.84</td>
</tr>
<tr>
<td>(n = 18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 5.2.* Distribution of *kanji* character corrected type token ratio per writing script over five different levels.

The one-way ANOVA test showed a statistically significant result with a large effect size:

\[ F(4, 141) = 29.70, \ p < .001, \eta^2 = 0.46 \]

As with KC RTTR, a post-hoc LSD test indicated that significant differences existed between all levels except between Level 7s and 9.

This measure showed a similar pattern to KC RTTR, confirming that the learners used a wider range of KCs in their writing as their level of proficiency
increased. However, once learners reached the higher-proficiency levels (i.e., Levels 7 and 9), they showed no further significant improvement.

5.2.1.3 Kanji word root type token ratio

Table 5.4 shows the descriptive statistics for mean scores and standard deviations of KW RTTR among students at five different proficiency levels. The box plots shown in Figure 5.3 illustrate the descriptive results. Again, a trend is shown in which learners increase substantially between Levels 1 and 3, then increase gradually between Levels 3 and 7, with a levelling off between Levels 7 and 9 (see Figure 5.3).

Table 5.4

Descriptive Statistics: Kanji Complexity—Kanji Word Root Type Token Ratio

<table>
<thead>
<tr>
<th>Level</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.41</td>
<td>1.12</td>
<td>.00</td>
<td>3.12</td>
</tr>
<tr>
<td>(n = 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3.62</td>
<td>0.78</td>
<td>2.12</td>
<td>5.28</td>
</tr>
<tr>
<td>(n = 40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4.14</td>
<td>0.78</td>
<td>2.62</td>
<td>5.86</td>
</tr>
<tr>
<td>(n = 45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>4.84</td>
<td>0.90</td>
<td>3.15</td>
<td>5.89</td>
</tr>
<tr>
<td>(n = 34)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>5.24</td>
<td>1.12</td>
<td>2.69</td>
<td>7.68</td>
</tr>
<tr>
<td>(n = 18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 5.3. Distribution of kanji word root type token ratio per writing script over five different levels.
The ANOVA test revealed a significant difference between the various proficiency levels, with a large effect size:

$$F(4, 141) = 36.91, \ p = < .001, \eta^2 = 0.51$$

Post-hoc LSD tests showed significant differences between all levels except between Levels 7 and 9.

Similar results from KC RTTR and KC CTTR were seen in this measure, suggesting that as the proficiency increases, learners use a wider variety of KCs as well as KWs. Nevertheless, the improvement was not significant between Levels 7 and 9.

**5.2.1.4 Kanji word corrected type token ratio**

The descriptive statistics in Table 5.5 and the box plots in Figure 5.4 indicate that the mean number for KW CTTR increased as the level of proficiency increased, showing a positive linear relationship. Similar to KW RTTR, there were overlaps between Levels 3 and 9 (see Figure 5.4).

Table 5.5

*Descriptive Statistics—Kanji complexity: Kanji Word Corrected Type Token Ratio*

<table>
<thead>
<tr>
<th>Level</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00</td>
<td>0.82</td>
<td>.00</td>
<td>2.26</td>
</tr>
<tr>
<td>(n = 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2.58</td>
<td>0.56</td>
<td>1.50</td>
<td>3.73</td>
</tr>
<tr>
<td>(n = 40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2.96</td>
<td>0.67</td>
<td>1.85</td>
<td>5.03</td>
</tr>
<tr>
<td>(n = 45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>3.43</td>
<td>0.64</td>
<td>2.23</td>
<td>4.16</td>
</tr>
<tr>
<td>(n = 34)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>3.70</td>
<td>0.83</td>
<td>1.91</td>
<td>5.43</td>
</tr>
<tr>
<td>(n = 18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The ANOVA test showed a significant difference between the various proficiency levels, with a large effect size:

$$F(4, 141) = 32.54, \ p < .001, \eta^2 = 0.48$$

The post-hoc test found a significant difference between all levels, except between Levels 7 and 9.

The results for KW CTTR also showed a positive linear relationship—a similar outcome to that of other three kanji complexity measures. The developmental changes suggest that as the level of the proficiency develops, learners are able to use a wider variety of KWS, although no significant improvement was observed between the higher levels (i.e., Levels 7 and 9).

5.2.2 *Kanji complexity results summary*

Overall, the findings for the measures for *kanji* complexity (i.e., KC RTTR, KC CTTR, KW RTTR, KW CTTR) showed a similar pattern. A positive linear relationship existed between the five different proficiency levels. However, significant improvement was not evident between Levels 7 and 9 for all measures discussed above. This suggests that once a learner reaches Level 7, it becomes difficult to include a wider variety of KCs and KWS in a written script. Nevertheless, it is important to note that the data
comprised written scripts from placement tests and learners had limited time in which to write. This limited time may have had a negative effect on improvement.

Although RTTR and CTTR for KCs and KWs have not been used as a measure to investigate L2 Japanese writing development in the past, these two measures were used by Park (2013) to analyse lexical complexity in English essays. Similar to Park’s study, the current study also suggests that these two measures can reveal developmental changes. As seen in the results discussed above, the ANOVA test indicated that statistically significant differences existed between the writing scripts from different proficiency levels learners—learners improved in each of these measures across each proficiency level until Level 7.

5.2.3 Syntactic complexity

The syntactic complexity measures included the ratio of te-form to words, the ratio of complex sentences to sentences, the ratio of clause conjunctions to words, and three different sentence-ending variables (propositional, epistemic and expressive). One-way ANOVA tests were conducted to determine if overall significantly different mean scores existed. Post-hoc LSD tests were then conducted to identify where the differences occurred. Data that did not meet the assumption of equality of variances were log transformed prior to analysis. The type of test performed for each measure is explained below, together with the results.

5.2.3.1 Ratio of te-form to words

Table 5.6 displays the descriptive statistics for the ratio- of te-form to words and Figure 5.5 mirrors these descriptive results. As seen in Table 5.6, the mean for the ratio of te-form to words increased gradually from Level 1 to 9, indicating a positive linear relationship. The increase from Level 1 to 3 was slightly greater compared to the other levels. Figure 5.5 shows that overlaps exist between Level 3 and Level 9 in the range of results, with ‘0’ being the minimum for all levels, except Level 9.
Table 5.6

Descriptive Statistics—Syntactic Complexity: Ratio of Te-form to Words

<table>
<thead>
<tr>
<th>Level</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.04</td>
<td>0.08</td>
<td>0.00</td>
<td>0.02</td>
</tr>
<tr>
<td>3</td>
<td>0.13</td>
<td>0.12</td>
<td>0.00</td>
<td>0.05</td>
</tr>
<tr>
<td>5</td>
<td>0.19</td>
<td>0.12</td>
<td>0.00</td>
<td>0.05</td>
</tr>
<tr>
<td>7</td>
<td>0.26</td>
<td>0.13</td>
<td>0.00</td>
<td>0.05</td>
</tr>
<tr>
<td>9</td>
<td>0.33</td>
<td>0.13</td>
<td>0.01</td>
<td>0.06</td>
</tr>
</tbody>
</table>

*Figure 5.5.* Distribution of the ratio of te-form to words per writing script over five different levels.

A one-way ANOVA test showed significant differences between the various proficiency levels, with a large effect size:

\[ F(4, 141) = 13.85, \ p < .001, \eta^2 = 0.28 \]

A more detailed analysis of the post-hoc test revealed significant differences between all levels, except between Level 7 and Level 9. In other words, no significant improvement was evident in the ratio of te-form for learners once learners had reached Level 7.

Although not great in number, te-form was also used as a measure in past studies that investigated Japanese L2 writing development (e.g., Kato & Kobayashi, 1985;
Tanaka, 1997). However, these studies only counted the frequency of te-form and did not perform any statistical tests. The fact that the current study used the ratio of te-form to words and conducted a statistical test has resulted in greater accuracy to confirm developmental changes.

5.2.3.2 Ratio of complex sentences to sentences

Table 5.7 displays the mean and standard deviation for the ratio of complex sentences to sentences, while the box plots in Figure 5.6 reflect the descriptive statistics. A gradual increase in the use of complex sentences is observed from Level 1 to 9. However, as seen in Figure 5.6, large overlaps in the range of responses were observed after the first level. Moreover, Level 5 contains a number of outliers.

Table 5.7

Descriptive Statistics—Syntactic Complexity: Ratio of Complex Sentences

<table>
<thead>
<tr>
<th>Level</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>(n = 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.11</td>
<td>0.14</td>
<td>0.00</td>
<td>0.67</td>
</tr>
<tr>
<td>(n = 40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.17</td>
<td>0.13</td>
<td>0.00</td>
<td>0.54</td>
</tr>
<tr>
<td>(n = 45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.26</td>
<td>0.15</td>
<td>0.06</td>
<td>0.70</td>
</tr>
<tr>
<td>(n = 34)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.31</td>
<td>0.15</td>
<td>0.07</td>
<td>0.53</td>
</tr>
<tr>
<td>(n = 18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 5.6. Distribution of ratio of complex sentences to sentences per writing script over five different levels.

The assumption of equality of variances was not satisfied for this measure; as such, a log transformation was first conducted. Then, to determine whether any statistically significant differences existed between the proficiency levels, a one-way ANOVA test was performed. The ANOVA test revealed significant differences between the five levels, with a large effect size:

$$F(4, 141) = 13.19, p < .001, \eta^2 = 0.27$$

A post-hoc LSD test was performed to determine where the exact differences between the proficiency levels were located. The post-hoc LSD test indicated that significant differences existed between all levels, except between Levels 3 and 5 and between Levels 7 and 9.

The results show that at Level 1, none of the learners could write complex sentences. This is somewhat predictable, as these participants are at the beginning stage of learning and can only write simple sentences. A proficiency increase was associated with a statistically significant increase in the ratio of complex sentences between Levels 1 and 3, then learners levelled off between Levels 3 and 5, showing no significantly different increase in the ratio. Another significant increase was then observed between Levels 5 and 7, but this was followed by a levelling off between Levels 7 and 9. In other words, instead of showing steady increase between the levels, the learners levelled off
and showed a slow increase at wider intervals. This suggests that it may take some time for learners to develop their ability to write complex sentences.

### 5.2.3.3 Ratio of clause conjunctions to words

Table 5.8 shows the descriptive statistics of the ratio of clause conjunctions to words. Figure 5.7 reflects the descriptive statistics shown in Table 5.8. As seen in Table 5.8 and Figure 5.7, as the level of proficiency developed, the ratio of clause conjunctions increased, revealing a positive linear relationship.

Table 5.8

<table>
<thead>
<tr>
<th>Level</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (n = 9)</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.04</td>
</tr>
<tr>
<td>3 (n = 40)</td>
<td>0.02</td>
<td>0.02</td>
<td>0.00</td>
<td>0.08</td>
</tr>
<tr>
<td>5 (n = 45)</td>
<td>0.03</td>
<td>0.01</td>
<td>0.00</td>
<td>0.06</td>
</tr>
<tr>
<td>7 (n = 34)</td>
<td>0.04</td>
<td>0.02</td>
<td>0.01</td>
<td>0.09</td>
</tr>
<tr>
<td>9 (n = 18)</td>
<td>0.06</td>
<td>0.03</td>
<td>0.04</td>
<td>0.16</td>
</tr>
</tbody>
</table>

*Figure 5.7. Distribution of ratio of clause conjunctions per writing script over five different levels.*
The one-way ANOVA test revealed statistically significant differences between the proficiency levels, with a small effect size:

\[
F(4, 141) = 23.06, \ p < .001, \ \eta^2 = 0.04
\]

A post-hoc LSD test was conducted to examine the differences precisely. The test revealed significant differences across all levels.

The results indicate that the ratio of clause conjunctions to words showed statistically significant improvement between each level. In other words, as learners’ proficiency improved, the use of clause conjunctions also increased. However, the effect size was small, which indicates that the improvement between levels was not large. The results in this current study contradict Tanaka’s (1997) findings; his study found a relatively large improvement in the participants’ use of clause conjunctions. This may be due to two reasons. First, as Tanaka (1997) counted the frequency of occurrence of clause conjunctions in the script, instead of looking at the ratio for each measure in relation to words (as this study does). Using ratios in this way would provide a more accurate result. Another possible reason may be that Tanaka’s study had a very small sample size (\(N = 10\)).

5.2.3.4 Ratio of propositional endings to words

Sentence-ending variables are useful in examining naturalness in Japanese writing (Tanaka, 1997). As discussed in Chapter 3, Section 3.2.1.3.4, a number of sentence-ending variables exist. Kato and Kobayashi (1985) compared the use of various sentence endings between non-native and native speakers. In relation to propositional endings, non-native speakers tended to use more propositional endings than did native speakers. Further, the findings also suggested that as native speakers mature, they tend to use fewer propositional endings. Hence, it can be said that if the ratio of propositional endings decreases as the level of proficiency increases, then
naturalness in learners’ writing becomes evident. This indicates an improvement in their writing.

The descriptive statistics in Table 5.9 and the box plots in Figure 5.8 show that the mean number for ratio of propositional endings decreased as the level of proficiency increased; however, the box plots indicate that the median did not decrease across each level. A decrease in ratio was the largest between Level 1 and 3. In addition, the range of ratios substantially overlapped between each level. A large number of outliers were also observed for Levels 1, 3, 5 and 7.

Table 5.9

*Descriptive Statistics—Syntactic Complexity: Ratio of Propositional Endings*

<table>
<thead>
<tr>
<th>Level</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.13</td>
<td>0.67</td>
<td>0.00</td>
<td>0.20</td>
</tr>
<tr>
<td>(n = 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.04</td>
<td>0.03</td>
<td>0.00</td>
<td>0.12</td>
</tr>
<tr>
<td>(n = 40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.03</td>
<td>0.02</td>
<td>0.00</td>
<td>0.11</td>
</tr>
<tr>
<td>(n = 45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.02</td>
<td>0.02</td>
<td>0.00</td>
<td>0.08</td>
</tr>
<tr>
<td>(n = 34)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.02</td>
<td>0.02</td>
<td>0.00</td>
<td>0.08</td>
</tr>
<tr>
<td>(n = 18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 5.8.* Distribution of ratio of propositional endings to words per writing script over five different levels.
As the assumption of equality of variance did not hold for this measure, a log transformation was used. Then, to examine whether any significant differences existed between the difference proficiency levels, a one-way ANOVA test was performed with the transformed data. The test revealed statistically significant differences between the proficiency levels, with a large effect size:

\[ F(4, 141) = 28.30, \ p < .001, \eta^2 = .45 \]

The post-hoc LSD test result revealed that significant differences were only apparent between Level 1 and all other levels; however, once learners reached Level 3, on average they did not improve (i.e., decrease) in their use of propositional endings.

The results indicate that the ratio of propositional endings was high at Level 1 and decreased considerably in Level 3; from Level 3, the decrease was not substantial.

5.2.3.5 **Ratio of epistemic endings to words**

The descriptive statistics for the ratio of epistemic endings among students at each proficiency level are provided in Table 5.10. Box plots illustrating this relationship are also shown in Figure 5.9. They indicate that after the first level, ratios were largely overlapping, as observed by the same mean score for each level, and the larger number of outliers in the data.

Differences between Level 1 and other levels were observed; however, epistemic endings were absent among scripts for all students in Level 1. The mean number for students in all other levels was 0.01.
Table 5.1

*Descriptive Statistics—Syntactic Complexity: Ratio of Epistemic Endings*

<table>
<thead>
<tr>
<th>Level</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>(n = 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.01</td>
<td>0.02</td>
<td>0.00</td>
<td>0.09</td>
</tr>
<tr>
<td>(n = 40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
<td>0.04</td>
</tr>
<tr>
<td>(n = 45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
<td>0.04</td>
</tr>
<tr>
<td>(n = 34)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.01</td>
<td>0.02</td>
<td>0.00</td>
<td>0.06</td>
</tr>
<tr>
<td>(n = 18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 5.9.* Distribution of ratio of epistemic endings to words per writing script over five different levels.

As the assumption of equality of variance was not satisfied in this case, a log transformation was used. The ANOVA test confirmed that statistically significant differences existed between proficiency levels, with a medium effect size:

\[ F(4, 141) = 3.22, \ p = 0.015, \ \eta^2 = 0.09 \]

In order to determine exactly where the significant differences were located between the proficiency levels, a post-hoc test was performed with the transformed data. The test revealed a significant difference in ratios between Level 1 and all other levels.
and between Levels 3 and 9. However, all other comparisons exhibited no statistically significant increases in the mean ratio of epistemic endings to words.

The results indicate that although an observed increase in the ratio of epistemic endings to words was apparent, statistically significant differences were between Level 1 and all other levels and between Levels 3 and 9. The fact that none of the Level 1 learners used epistemic endings may suggest that this measure is not appropriate when observing beginner-level Japanese L2 learners.

5.2.3.6 Ratio of expressive endings to words

The descriptive statistics of the ratio of expressive endings to words is shown in Table 5.11. Box plots in Figure 5.10 illustrate the descriptive statistics. They indicate that the mean number of the ratio of expressive endings did not increase consistently as the learners’ proficiency developed. An increase from Level 1 to 3 was noticeable, and this remained constant from Level 3 to 5. Although a slight drop in the ratio existed at Level 7, it increased between Levels 7 and 9. Moreover very little difference was evident in the mean ratios across the five different levels.

Table 5.11

Descriptive Statistics—Syntactic Complexity: Ratio of Expressive Endings

<table>
<thead>
<tr>
<th>Level</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>(n = 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.02</td>
<td>0.01</td>
<td>0.00</td>
<td>0.06</td>
</tr>
<tr>
<td>(n = 40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.02</td>
<td>0.01</td>
<td>0.00</td>
<td>0.07</td>
</tr>
<tr>
<td>(n = 45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
<td>0.05</td>
</tr>
<tr>
<td>(n = 34)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.02</td>
<td>0.02</td>
<td>0.00</td>
<td>0.05</td>
</tr>
<tr>
<td>(n = 18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 5.10. Distribution of ratio of expressive endings to words per writing script over five different levels.

The equality of variance assumption was not violated in this case; hence, a one-way ANOVA test was conducted to examine whether any statistically significant differences existed between the different proficiency levels. The ANOVA test revealed significant differences between the various proficiency levels, with a small effect size:

$$F(4, 141) = 4.33, p = .002, \eta^2 = 0.01$$

In order to determine where the exact differences in the data were located, a post-hoc LSD test was analysed. This determined that significant differences were only observed between Level 1 and all other levels. No other comparisons exhibited statistically significant differences.

The results indicate that the ratio of expressive endings did not have a positive linear relationship. The only statistically significant differences were between Level 1 and all other levels. Further, learners in Level 1 did not use any expressive endings and learners in other levels only used small percentage of expressive endings in their writing. Considering that this was a small effect size, even though significant differences existed between Level 1 and all other levels, the improvement for this measure between Level 1 and other levels was very small.
Tanaka (1997) indicated that the percentage of expressive endings used by learners after in-country experience was 16% as to 2.4% higher than before their in-country experience. One reason for the greater increase in the percentage seen in Tanaka’s study, compared to the current study, may be the difference in the writing task topic. Another factor could be the influence of a learner’s spoken language after the in-country experience. Tanaka (1997) indicated that many learners used expressions such as ‘dayo’ (‘you know’, ‘anyway’) and ‘dane’ (tag question, ‘you know’) after their in-country experience. However, in his study, statistical analysis was not undertaken to show whether a significant difference existed between learners’ writing before and after their in-country experience. As such, the results of Tanaka’s study may not be reliable.

5.2.4 Syntactic complexity results summary

The results for syntactic complexity were mixed. The ratio of clause conjunctions was one of the only measures that showed significant differences between all levels. However, it is worth noting that the effect size for the ratio of clause conjunctions was small, indicating that the improvement was subtle. Developmental changes were also observed for the ratio of te-form and the ratio of complex sentences. Significant differences were not apparent in the higher-proficiency levels for the ratio of te-form. Similarly, developmental changes were not significant between Levels 3 and 5 and between Levels 7 and 9 for the ratio of complex sentences.

The three sentence-ending variables (propositional, epistemic and expressive endings) showed fewer developmental changes compared to the three syntactic measures discussed above. The ratio of sentence-ending variables only showed significant differences between Level 1 and other levels (epistemic endings also showed a significant difference between Levels 3 and 9).

In comparison to kanji complexity, developmental changes for syntactic complexity were less significant between the proficiency levels. The results suggest that
syntactic complexity may be difficult to improve, especially for those learning a second language in a FL context, as with participants in the current study. Although detailed background information was unavailable in the current study, the participants in this study should be considered FL learners. This is because even if some have been to Japan in the past (i.e., during secondary school), as the majority of their time was spent in Australia, they cannot be considered equivalent to SL learners. However, it is also important to note that use of the terminology ‘FL learners’ and ‘SL learners’ has not been clear in recent years. According to Ortega (2003), students may have difficulty developing syntactic complexity in FL contexts. Ortega’s (2003) syntactic complexity study suggests that FL learners may develop less fully than do L2 learners.

5.2.5 Accuracy

Chosen accuracy measures include the ratio of error-free sentences, the ratio of EFCs to clauses, and words per error-free sentences. For all accuracy measures, the assumption of equality of variances was not violated. A one-way ANOVA test was conducted to determine if a significant difference existed between the different proficiency levels. Where the ANOVA test showed significant differences, a post-hoc LSD test was also conducted to determine where the significant differences were located between the levels.

5.2.5.1 Ratio of error-free sentences to sentences

The descriptive statistics in Table 5.12 and the box plots in Figure 5.11 below indicate that as the level of proficiency increased, the ratio of error-free sentences also increased. From Level 1 to 7, a steady increase was observed and the increase was greater between Levels 7 and 9. As seen in Figure 5.11, the range of ratios overlapped between all levels.
Table 5.12

Descriptive Statistics—Accuracy: Ratio of Error-free Sentences

<table>
<thead>
<tr>
<th>Level</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.27</td>
<td>0.13</td>
<td>0.10</td>
<td>0.50</td>
</tr>
<tr>
<td>3</td>
<td>0.29</td>
<td>0.15</td>
<td>0.00</td>
<td>0.75</td>
</tr>
<tr>
<td>5</td>
<td>0.35</td>
<td>0.17</td>
<td>0.08</td>
<td>0.69</td>
</tr>
<tr>
<td>7</td>
<td>0.42</td>
<td>0.15</td>
<td>0.05</td>
<td>0.67</td>
</tr>
<tr>
<td>9</td>
<td>0.60</td>
<td>0.20</td>
<td>0.24</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Figure 5.11. Distribution of ratio of error-free sentences per writing script over five different levels.

Results from the ANOVA test indicate that statistically significant differences exist between the proficiency levels, with a large effect size:

\[ F(4, 141) = 13.38, \ p < .001, \eta^2 = 0.28 \]

To determine precisely where the significant differences between the proficiency levels were located, a post-hoc LSD test was analysed. The test revealed that significant differences existed between Level 7 and all other levels, as well as Level 9 and all other levels.
The results indicate that even though an observed increase in the ratio of error-free sentences to sentences existed, in the levels with statistically significant differences these were between the higher levels (i.e., Levels 7 and 9) and all other levels. This suggests that unless a learner reaches Level 7, improvement for this measure is not significant.

The past literature that employed the ratio of error-free sentences (e.g., Tanaka, 1997; Tapia, 1993) showed this was not a significant measure, which contradicts the current study’s findings. The learners’ proficiency levels in Tanaka’s study were not advanced; this may be why the measure showed no development in his study.

5.2.5.2 Ratio of error-free clauses to clauses

As seen in descriptive statistics in Table 5.13 and the box plots in Figure 5.12, the ratio of EFCs showed a positive linear relationship. As the learners’ level of proficiency increased, an increase also occurred in the ratio of EFCs. The increase from Level 1 to 3 was the greatest.

Table 5.13

Descriptive Statistics—Accuracy: Ratio of Error-free Clauses

<table>
<thead>
<tr>
<th>Level</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.15</td>
<td>0.34</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>(n = 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.36</td>
<td>0.29</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>(n = 40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.51</td>
<td>0.20</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>(n = 45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.59</td>
<td>0.19</td>
<td>0.29</td>
<td>1.00</td>
</tr>
<tr>
<td>(n = 34)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.72</td>
<td>0.17</td>
<td>0.40</td>
<td>1.00</td>
</tr>
<tr>
<td>(n = 18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The ANOVA test revealed that statistically significant differences existed between the proficiency levels, with a large effect size:

$$F(4, 141) = 14.24, p < .001, \eta^2 = 0.29$$

The post-hoc LSD test revealed that significant differences existed between all levels, except between Levels 5 and 7.

The results indicate that learners were able to write more EFCs in a writing script as their proficiency developed, although the development was not significant between Levels 5 and 7. The findings of this study were similar to those of Ishikawa (1995), Cumming et al. (2006) and Knoch, Macqueen et al. (2014), who all found that higher-proficiency learners produced more EFCs.

5.2.5.3 Words per error-free sentence

The descriptive statistics in Table 5.14 and the box plots in Figure 5.13 indicate that words per error-free sentences gradually increased from Level 1 to 9. However, as seen in Figure 5.13, overlaps occurred between all levels. Moreover, the data contained a few outliers.
Table 5.14

Descriptive Statistics—Accuracy: Words per Error-free Sentences

<table>
<thead>
<tr>
<th>Level</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.22</td>
<td>1.55</td>
<td>2.00</td>
<td>7.00</td>
</tr>
<tr>
<td>(n = 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>5.80</td>
<td>1.78</td>
<td>0.00</td>
<td>9.30</td>
</tr>
<tr>
<td>(n = 40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6.66</td>
<td>2.24</td>
<td>2.00</td>
<td>11.70</td>
</tr>
<tr>
<td>(n = 45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>8.00</td>
<td>2.59</td>
<td>2.30</td>
<td>14.80</td>
</tr>
<tr>
<td>(n = 34)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>9.76</td>
<td>2.22</td>
<td>7.00</td>
<td>16.30</td>
</tr>
<tr>
<td>(n = 18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 5.13.* Distribution of words per error-free sentences per writing script over five different levels.

The results from the ANOVA test show that statistically significant differences exist between the proficiency levels, with a large effect size:

\[
F (4, 141) = 13.78, \ p < .001, \ \eta^2 = 0.28
\]

The post-hoc LSD test revealed that significant differences existed between Level 7 and all other levels, and between Level 9 and all other levels.

The results show that although an increase in words per error-free sentence was apparent, the levels with statistically significant differences only occurred between the upper levels and all other levels. This result was similar to that observed for the ratio of
error-free sentences to sentences. Thus, improvement in this measure becomes significant when the learner reaches Level 7. The measure of words per error-free sentence was also used in Tanka’s (1997) study, where the learners showed improvement for this measure between the pre- and post-tests. However (as mentioned earlier), despite the similar results, these studies may not be directly comparable due to different methodologies.

5.2.6 Accuracy results summary

Three accuracy measures were examined. The findings revealed that the ratio of EFCs showed a steady development as the level of proficiency increased, although it did not increase significantly between Levels 5 and 7. The developmental changes for the ratio of error-free sentences and words per error-free sentences were only significant between the upper proficiency levels (Levels 7 and 9) and all other levels. This suggests that with these two measures, the developmental differences are only apparent once the learner reaches Level 7. In other words, higher-proficiency level learners tend to write more accurately. This outcome was also evident in past studies (e.g., Cumming et al., 2006; Knoch, Macqueen et al., 2014)

5.2.7 Fluency

The fluency measures included the number of words, the number of sentences, the number of KCs, the number of KWs and the number of clauses. Similar to the syntactic complexity measures, the assumption of equality of variances was not satisfied with some fluency measures. As such, for those measures, a log transformation was first conducted and then a one-way ANOVA test was performed. A post-hoc LSD test was also undertaken with the transformed data to determine where the significant differences were present in the data. For each measure, the type of test performed is explained below, together with the results.
5.2.7.1 Number of words

As seen from the descriptive statistics in Table 5.15 and the box plots in Figure 5.14, a positive linear relationship exists between the number of words and the learners’ proficiency level. In other words, as the proficiency level increased, learners were able to write more words in the writing script. The largest improvement was seen between Levels 1 and 3. Large overlaps were present between Levels 3 and 9 (see Figure 5.14).

Table 5.15

**Descriptive Statistics—Average Number of Words per Script**

<table>
<thead>
<tr>
<th>Level</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>32.11</td>
<td>12.04</td>
<td>16.00</td>
<td>48.00</td>
</tr>
<tr>
<td>(n = 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>134.58</td>
<td>59.08</td>
<td>31.00</td>
<td>314.00</td>
</tr>
<tr>
<td>(n = 40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>179.42</td>
<td>49.85</td>
<td>75.00</td>
<td>296.00</td>
</tr>
<tr>
<td>(n = 45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>200.41</td>
<td>48.50</td>
<td>125.00</td>
<td>318.00</td>
</tr>
<tr>
<td>(n = 34)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>247.33</td>
<td>65.58</td>
<td>151.00</td>
<td>382.00</td>
</tr>
<tr>
<td>(n = 18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 5.14*. Distribution of number of words per writing script over five different levels.

The ANOVA analysis revealed statistically significant differences between the different proficiency levels and a large effect size:

\[
F(4, 141) = 32.31, p < .001, \eta^2 = 0.48
\]
The post-hoc LSD test showed that significant differences existed between all levels, except between Levels 5 and 7.

The results reveal an increase in the number of words, suggesting that as the learners’ proficiency developed, the learners wrote more words in the script. The only levels that did not show statistically significant differences were those between Levels 5 and 7. The findings on the number of words in the current study are similar to those of Tanaka’s (1997) study, which also showed that the number of words improved from pre- to post-test. In fact, this measure (number of words) was the area in which students made the most improvement in Tanaka’s study.

5.2.7.2 Number of sentences

The descriptive statistics in Table 5.16 and the box plots in Figure 5.15 indicate that as the level of proficiency increased, the mean number of sentences also increased. The difference between Levels 1 and 3 was the largest, suggesting that the greatest development took place between these levels. Although development occurred between other levels (e.g., Levels 3 and 5), the differences were not as substantial. Further, overlaps existed between Levels 3 to 9.

Table 5.16

Descriptive Statistics—Average Number of Sentences per Script

<table>
<thead>
<tr>
<th>Level</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.22</td>
<td>2.43</td>
<td>3.00</td>
<td>10.00</td>
</tr>
<tr>
<td>(n = 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>13.60</td>
<td>4.96</td>
<td>4.00</td>
<td>25.00</td>
</tr>
<tr>
<td>(n = 40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>15.71</td>
<td>4.50</td>
<td>7.00</td>
<td>24.00</td>
</tr>
<tr>
<td>(n = 45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>16.97</td>
<td>4.80</td>
<td>10.00</td>
<td>32.00</td>
</tr>
<tr>
<td>(n = 34)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>19.67</td>
<td>5.65</td>
<td>10.00</td>
<td>29.00</td>
</tr>
<tr>
<td>(n = 18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 5.15. Distribution of number of sentences per writing script over five different levels.

The ANOVA test showed a statistically significant result with a large effect size:

$$F(4, 141) = 14.38, \ p < .001, \ \eta^2 = 0.29$$

The post-hoc LSD tests revealed that significant differences existed between all levels, except between Levels 5 and 7 and between Levels 7 and 9.

The findings above show that development occurred between the various proficiency levels, although no significant development was apparent between Levels 5 and 7 and between Levels 7 and 9. This indicates that proficiency is associated with improvement in this measure only in the lower levels. Other studies (e.g., Iwashita & Sekiguchi, 2009; Tanaka, 1997) that examined the number of sentences for their fluency measures also found improvement in participants’ number of sentences. Participants in Iwashita and Sekiguchi’s (2009) and Tanaka’s (1997) study were not advanced learners. Similarly, the levels with significant differences in this study were not in the upper range, which concurs with the findings from other studies.

5.2.7.3 Number of kanji characters

The descriptive statistics in Table 5.17 and the box plots in Figure 5.16 present the number of KCs for each level. They show that as proficiency levels increased, the number of KCs used also increased. The difference between Levels 1 and 3 was the greatest.
Table 5.17

Descriptive Statistics—Average Number of Kanji Characters per Script

<table>
<thead>
<tr>
<th>Level</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (n = 9)</td>
<td>11.00</td>
<td>15.48</td>
<td>.00</td>
<td>46.00</td>
</tr>
<tr>
<td>3 (n = 40)</td>
<td>49.20</td>
<td>24.11</td>
<td>19.00</td>
<td>134.00</td>
</tr>
<tr>
<td>5 (n = 45)</td>
<td>59.20</td>
<td>23.50</td>
<td>12.00</td>
<td>115.00</td>
</tr>
<tr>
<td>7 (n = 34)</td>
<td>77.91</td>
<td>30.69</td>
<td>27.00</td>
<td>173.00</td>
</tr>
<tr>
<td>9 (n = 18)</td>
<td>95.78</td>
<td>43.23</td>
<td>34.00</td>
<td>187.00</td>
</tr>
</tbody>
</table>

Figure 5.16. Distribution of number of kanji characters per writing script over five different levels.

As the assumptions of equality of variance were not satisfied in this case, a log transformation was used. A one-way ANOVA test was then performed with the transformed data. The test revealed a statistically significant result, with a large effect size:

\[ F(4, 141) = 45.98, \ p < .001, \ \eta^2 = 0.35 \]

A post-hoc LSD test was conducted with the transformed data to identify the significant differences between levels. The post-hoc LSD test indicated that significant differences existed between each level, except between Levels 3 and 5.
The above result suggests that the number of KCs learners can write in Level 1 is minimal. However, from Level 1 to 3, the increase in the number of KCs used is substantial. This suggests that learners are able to acquire a number of KCs during that period. From Level 3 to 9, a steady increase was observed, although this increase was not significant between Levels 3 and 5.

**5.2.7.4 Number of kanji words**

As seen in the descriptive statistics in Table 5.18 and the box plots in Figure 5.17, a positive linear relationship was observed between the number of KWs and the level of proficiency. In other words, as the level of proficiency increased, the number of KWs also increased. Similar to the number of KCs, the greatest increase was seen from Level 1 to 3. In addition, overlaps were observed from Level 3 onwards.

Table 5.18

*Descriptive Statistics—Average Number of Kanji Words per Script*

<table>
<thead>
<tr>
<th>Level</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.67</td>
<td>7.37</td>
<td>.00</td>
<td>22.00</td>
</tr>
<tr>
<td>(n = 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>31.48</td>
<td>14.72</td>
<td>11.00</td>
<td>76.00</td>
</tr>
<tr>
<td>(n = 40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>39.67</td>
<td>13.74</td>
<td>9.00</td>
<td>68.00</td>
</tr>
<tr>
<td>(n = 45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>51.32</td>
<td>20.28</td>
<td>17.00</td>
<td>116.00</td>
</tr>
<tr>
<td>(n = 34)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>61.72</td>
<td>26.11</td>
<td>27.00</td>
<td>116.00</td>
</tr>
<tr>
<td>(n = 18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 5.17. Distribution of number of kanji words per writing script over five different levels.

As the assumption of variance was not satisfied in this case, a log transformation was applied and an ANOVA test with the transformed data showed a statistically significant result, with a large effect size:

\[ F(4, 141) = 21.88, \ p < .001, \ \eta^2 = 0.38 \]

To determine where the significant differences were located, a post-hoc LSD test was performed. The test revealed that significant differences existed between all levels.

The results above indicate that the increase in the number of KWs was observed with statistically significant differences between all five proficiency levels. This suggests that as the learner’s proficiency developed, not only can they acquire a number of KWs, but they can also retain and use them.

5.2.7.5 Number of clauses

It is clear from the descriptive statistics in Table 5.19 and the box plots in Figure 5.18 that as the level of proficiency increased, the number of clauses increased. This indicates a positive linear relationship. Overall, a steady increase existed from Level 1 to 9. However, the increase in the lower proficiency levels (e.g., between Levels 1 and 3) was slightly higher compared to the higher-proficiency levels (e.g., between Levels 7 and 9). Overlaps existed between Levels 5 and 9.
Table 5.19

Descriptive Statistics—Average Number of Clauses per Script

<table>
<thead>
<tr>
<th>Level</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00</td>
<td>1.12</td>
<td>.00</td>
<td>3.00</td>
</tr>
<tr>
<td>(n = 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>5.23</td>
<td>3.03</td>
<td>.00</td>
<td>13.00</td>
</tr>
<tr>
<td>(n = 40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>10.73</td>
<td>3.96</td>
<td>3.00</td>
<td>21.00</td>
</tr>
<tr>
<td>(n = 45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>13.32</td>
<td>4.58</td>
<td>5.00</td>
<td>23.00</td>
</tr>
<tr>
<td>(n = 34)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>16.89</td>
<td>4.99</td>
<td>8.00</td>
<td>25.00</td>
</tr>
<tr>
<td>(n = 18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 5.18.* Distribution of number of clauses per writing script over five different levels.

As the assumption of equality of variance was not satisfied, a log transformation was performed. A one-way ANOVA test was then conducted with the transformed data. The ANOVA analysis with the transformed data revealed statistical differences between the different levels of proficiency, with a large effect size:

\[ F(4,141) = 46.92, p < .001, \eta^2 = 0.57 \]

To determine exactly where the significant differences were located, a post-hoc LSD test was performed with the transformed data. This revealed significant differences between all levels.
The results indicate that a positive linear relationship exists between the number of clauses and the proficiency levels, showing significant differences between all levels. The fact that there were significant differences between all levels suggests that this measure can be used when observing developmental changes for any proficiency level (i.e., beginner, intermediate or advanced).

5.2.8 Fluency results summary

Overall, the findings for the fluency measures showed developmental changes between the five different proficiency levels, although some measures (e.g., the number of KCs) showed no significant developmental changes between Levels 3 and 5. However, compared to other measures, fluency measures revealed differences from one level to another. It appears that the fluency measure is one of the most likely to improve, compared to other measures (e.g., syntactic complexity, accuracy). For example, in Tanaka’s (1997) study, students improved the most in fluency. In another study conducted by Knoch et al. (2015) fluency was the only measure that showed improvement in ESL students’ writing after three years.

5.3 Conclusion

This chapter has detailed the results based on an analysis of 146 writing scripts. These scripts came from the writing task of a placement test administered in Semester 1, 2013 in a large metropolitan Australian university. Box plots provided a clear overview of the distribution, plotting each measure against the different proficiency levels. Inferential statistics were also shown for each measure.

Four different writing features were examined: kanji complexity, syntactic complexity, accuracy and fluency, with each containing a number of measures. Overall, the measures investigated for kanji complexity and fluency exhibited more developmental changes compared to syntactic complexity and accuracy. These results suggest that syntactic complexity and accuracy measures may be more difficult to
examine regarding developmental changes, depending on the learners’ proficiency level. For example, two of the accuracy measures (the ratio of error-free sentences and words per error-free sentences) showed that improvements were only significant between the upper proficiency levels (Levels 7 and 9) and all other levels. This suggests that developmental differences can only be seen once a learner reaches Level 7.

As discussed above, the ANOVA tests revealed the measures investigated in Study 1 showed significant differences between the overall group.

Nevertheless, the post-hoc LSD tests revealed that not every measure showed significant differences between every level. Most measures showed developmental changes in the lower- and intermediate-proficiency level ranges, except for a few measures of accuracy. Hence, all the measures investigated in Study 1 were considered for use in Study 2 to analyse writing development, considering that participants in Study 2 were lower-intermediate learners.
Chapter 6: Study 2 Methodology

Study 2 of this research project employed a longitudinal mixed-methods design to expand our understanding of L2 writing development in an L2 Japanese context. The study used a pre- and post-test design for one semester (short term) as well as one year (long term) to examine writing development. The study also investigated the differences in writing development between character-based learners and non-character-based learners. Further, drawing on the data attained through student and teacher interviews, Study 2 examined both students’ and teachers’ perspectives on writing development.

Study 2 was conducted at the same university as Study 1. This chapter begins with an overview of the context of the Japanese course offered by the university. It then describes the study design, as well as the participant details. This is followed by a description of the instruments and procedures used for data collection. The chapter ends with a description of the data analysis procedures.

6.1 Context

Japanese is offered as a major in the Bachelor of Arts over three years at the university, but students can also take Japanese as an elective. Japanese 1 to Japanese 8 classes are offered as semester-long subjects in a progressive sequence; Japanese 1 being the beginner-level subject and Japanese 8 the advanced level subject. Study 2 aimed to investigate the writing development of lower-intermediate level learners, the equivalent of Japanese 3 and Japanese 4. Accordingly, Study 2 was conducted in these two subjects.

6.1.1 Japanese 3 (Semester 1)

Japanese 3 caters for students at the lower-intermediate level. In Semester 1, 2014, when the study was conducted, Japanese 3 was taught by one lecturer (coordinator) and four sessional tutors who were native Japanese speakers. Two-hundred-and-four students were enrolled in this subject.
The contact hours consisted of a 1-hour lecture and two 1.5-hour seminars per week (i.e., 3 hours of seminar time per week). The subject’s objective was for students to learn around 100 new KCs, report facts and experiences in Japanese and develop some cultural understanding.

The lecture (attended by all enrolled students) conveyed information about Japanese culture; for example, imparting an understanding of Japanese social norms and customs about gift giving and receiving. In the seminars (consisting of eight parallel streams for both Seminar 1 and Seminar 2), students were taught new grammar and new *kanji* every week. They were also involved in various activities, including peer interaction to discuss questions in the textbook relating to cultural awareness; for example, on how to exchange gifts or make suggestions.

Several assessments were issued throughout the semester (see Table 6.1), but the course also included weekly *kanji* dictation (sentence dictation including 10 new *kanji*) and weekly reflective writing tasks (3 sentences at the end of class), which were not assessed. These tasks were completed during seminars.

Table 6.1

*Japanese 3 Assessments and Requirements*

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Three-stage small cultural group projects (done in groups of three) Due throughout semester</td>
<td>Purpose: prepare students for their individual assignments. The group project consisted of three stages: 1) Write an itinerary (300 words in Japanese) from Narita airport to Amanohashidate, a tourist destination in Japan. 2) Write a discussion piece on the destination (approximately 600 words in English). 3) Write an annotated bibliography (600–800 words in English). Produce a tourist webpage in Japanese</td>
</tr>
<tr>
<td>Individual written work Due end of semester</td>
<td>Produce a tourist webpage in Japanese (600 words)</td>
</tr>
<tr>
<td>Two oral assessments 1) Due mid-semester 2) Due at the end of semester</td>
<td>1) Create a 5-minute video in a group (same group as the cultural project group) introducing a place in Melbourne 2) A 5-minute speaking test where students answer questions on several randomly chosen topics covered in</td>
</tr>
</tbody>
</table>
2-hour written examination
End of semester

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-stage small cultural group projects (done in groups of three)</td>
<td>Purpose: prepare students for their individual assignments. The group project consisted of two stages: 1) Write about how the group chose a particular topic (600 words in English). 2) Write an annotated bibliography (600 words in English).</td>
</tr>
</tbody>
</table>

6.1.2 Japanese 4 (Semester 2)

In Japanese 4, the contact hours were the same as for Japanese 3 (a 1-hour lecture and two 1.5-hour seminars per week). The subject aimed to introduce an additional 100 new KCs, teach different expressions (honorific, humble and polite) and provide cultural awareness for intercultural communication in different settings in Japan.

In 2014, Japanese 4 ran in Semester 2. It was taught by the same lecturer (coordinator) who had taught Japanese 3, along with two sessional tutors who had also taught Japanese 3. There were 125 students enrolled in the subject (forming five parallel stream for both Seminar 1 and Seminar 2).

Apart from the assessments shown in Table 6.2, students also had weekly kanji dictation (sentence dictation including 10 new kanji) and reflective writing tasks (3 sentences at the end of the seminars), which were not assessed. For the weekly reflective writing tasks, in some weeks students were required to undertake a peer review.

Table 6.2

Japanese 4 Assessments and Requirements

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peer review assignment</td>
<td>Peer review assignment consisting of a two-step</td>
</tr>
</tbody>
</table>
**Assessment** | **Requirement**
--- | ---
Due mid-semester process: 1) Submit a draft of their individual assignment to two other students 2) Give feedback on two other students’ assignments.

Individual written work A research task, where students could write on any topic covered in the lectures during the semester (600 words).

Due at the end of semester Two oral assessments 1) Create a 5-minute video in a group (same group as the cultural project group) using different expressions (honorific, humble and polite) in a setting of the students’ choice. 2) 5-minute speaking test where students answer questions on several random topics from the semester.

Due at the end of semester Two oral assessments 1) Due mid-semester 2) Due at the end of semester Two oral assessments 1) Due mid-semester 2) Due at the end of semester

2-hour written examination The 2-hour exam consisted of fill-in-the-blank questions, matching questions, multiple-choice questions and a series of three readings with short-answer questions at the end. In the exam, there was also a small writing task where students had to write a response to an email.

End of semester

---

**6.2 Study Design**

Study 2 was a one-year longitudinal study that aimed to investigate the individual writing development of lower-intermediate level learners. It consisted of two phases: one over a semester, and the other spanning a year. Ethics clearance was obtained prior to recruiting participants for this study. This section will describe the designs of both phases.

**6.2.1 Phase 1 design: one semester—Japanese 3**

Out of the 16 streams (eight streams in Seminar 1 and eight in Seminar 2), the researcher visited 11 to recruit participants. Forty-one participants were recruited between Weeks 3 and 6. Volunteer participants completed ethics forms (see Appendix F for the student consent form and Appendix G for the teacher consent form). At the time of recruitment, student participants also completed a background questionnaire (see Appendix B and the description in Section 6.3.2). During the semester, the researcher also attended and observed one lecture and one seminar every week (Week 3 to 10). The
seminar stream the researcher attended was taught by one tutor and three of the study participants were in that seminar stream.

The study used a pre- and post-test design (see Section 6.3.1). An in-class test conducted as part of the regular classroom activities during the first week of Japanese 3 was used as the pre-test. The post-test was conducted in week 12 outside class hours, but under the researcher’s supervision. Students were given the same amount of time (30 minutes) to complete the pre- and post-tests. Further, for both pre and post-tests participants were allowed to consult their dictionary but no other materials. After Japanese 3 concluded, 30-minute interviews were conducted individually with each of the 40 students (out of the 41 participants, one student did not participate in the interview) as well as two teachers (one sessional tutor and one lecturer). The schedule for the Phase 1 data collection activities outlined above is given in Table 6.3.

Table 6.3

Schedule for Japanese 3 Data Collection

<table>
<thead>
<tr>
<th>Research Activity</th>
<th>Japanese 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>Week 1</td>
</tr>
<tr>
<td>Participant recruitment</td>
<td>Week 3–Week 6</td>
</tr>
<tr>
<td>Background questionnaires</td>
<td>Week 3–Week 6</td>
</tr>
<tr>
<td>Lecture observation</td>
<td>Week 3–Week 10</td>
</tr>
<tr>
<td>Classroom observation</td>
<td>Week 3–Week 10</td>
</tr>
<tr>
<td>Post-test</td>
<td>Week 12</td>
</tr>
<tr>
<td>30-minute interviews</td>
<td>Week 14, Week 15 &amp; Week 16</td>
</tr>
<tr>
<td></td>
<td>(after conclusion of semester)</td>
</tr>
<tr>
<td>Collection of individual assignments</td>
<td>Week 16</td>
</tr>
</tbody>
</table>

Of the 41 participants who participated in Phase 1 of Study 2, 22 continued to study Japanese 4 (as Japanese is an elective subject, not all students continue studying it) and all agreed to participate in Phase 2 of Study 2.
6.2.2 Phase 2 design: one year (Japanese 3 & Japanese 4)

The 22 participants who continued to study Japanese 4 completed new ethics forms (see Appendix H). During Japanese 4, the researcher also attended and observed one lecture and one seminar-every week (Week 3 to 10). Similar to Phase 1, the seminar stream that the researcher observed was taught by one tutor and four of my participants attended that seminar stream.

In Week 13 of Japanese 4, participants were asked to do another post-test. The post-test in Japanese 4 was compared with the pre-test in Japanese 3 to examine development after one year (2 semesters) of language study. Although there were 41 participants from Japanese 3, only the pre-test from 22 participants were used for the one-year investigation. The post-test was conducted outside class hours, but under the researcher’s supervision and students were given the same time (30 minutes) to complete the post-test. As with Japanese 3, the participants were able to consult their dictionary.

After Japanese 4 concluded, a 30-minute interview was conducted with the course coordinator and each of the 22 students. The schedule for the Japanese 4 activities outlined is given in Table 6.4.

Table 6.4

Schedule for Japanese 4 Data Collection

<table>
<thead>
<tr>
<th>Research Activity</th>
<th>Japanese 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-test</td>
<td>Week 13</td>
</tr>
<tr>
<td>Lecture observation</td>
<td>Week 3–Week 10</td>
</tr>
<tr>
<td>Classroom observation</td>
<td>Week 3–Week 10</td>
</tr>
<tr>
<td>30-minute interviews</td>
<td>Week 16 &amp; Week 17 (after conclusion of semester)</td>
</tr>
<tr>
<td>Collection of individual assignments</td>
<td>Week 16</td>
</tr>
</tbody>
</table>

177
6.2.3 Participants

6.2.3.1 Student participants

As mentioned previously, the student participants in this study comprised 41 students from Japanese 3; 22 of the 41 students continued to Japanese 4. More female than male students were enrolled in Japanese 3 and 4 and this was also reflected in the study’s composition of participants. There were 28 female participants and 13 male participants from Japanese 3, and 15 female and 7 male participants from Japanese 4.

As many students take Japanese as an elective subject, participants came from different faculties (e.g., Arts and Science). They were enrolled in undergraduate or graduate courses. Thirty-nine of the 41 participants in Japanese 3 were undergraduate students, with the remaining two graduate students. In Japanese 4, 21 participants were undergraduates and one was a graduate student.

A background questionnaire was used to elicit the learners’ language background. Participants were divided into character-based learners and non-character-based learners. Learners familiar with Chinese characters (e.g., students from China, Hong Kong and Taiwan) were classified as character-based learners. Participants from Korea were classified as non-character-based learners (although some Koreans are familiar with kanji, the questionnaire revealed that participants in this study were not). Table 6.5 shows the number of participants of character-based language learners and non-character-based language learners for Japanese 3 and Japanese 4.

Table 6.5

<table>
<thead>
<tr>
<th>Participants</th>
<th>Japanese 3</th>
<th>Japanese 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Character-based learners</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>Non-character-based learners</td>
<td>25</td>
<td>14</td>
</tr>
<tr>
<td>Total</td>
<td>41</td>
<td>22</td>
</tr>
</tbody>
</table>
6.2.3.2 Teacher participants

This study had two teacher participants: one lecturer/coordinator (pseudonym: Hayashi-sensei) and one sessional teacher (pseudonym: Tanaka-sensei). Both participated in the interviews for Japanese 3 at the end of Semester 1, 2014. Only the lecturer, Hayashi-sensei, participated in the interview conducted at the end of Japanese 4 (end of Semester 2, 2014). In addition, a follow-up interview was conducted with Hayashi-sensei in early 2016.

Hayashi-sensei had been teaching Japanese since 1993 and has taught at both secondary school and university, from beginner to advanced levels. At this university (where the research was conducted), Hayashi-sensei has coordinated Japanese 3 and Japanese 4 since 2010. For Japanese 3 (Semester 1, 2014), Hayashi-sensei gave one lecture a week and taught five 1.5-hour seminars (i.e., five streams). In Japanese 4 (Semester 2, 2014) Hayashi-sensei again lectured once a week and taught six 1.5-hour seminars (i.e., six streams). Hayashi-sensei designed both the classroom activities and assessment tasks for both Japanese 3 and Japanese 4.

Tanaka-sensei had been teaching Japanese for about 20 years and has taught from beginner to advanced level. However, Japanese 3 (Semester 1, 2014) was the first semester in which Tanaka-sensei had taught at this university, as she had moved from another state (in Australia). Therefore, the teaching system was new for her in Japanese 3. Further, in Japanese 3 she only taught one seminar stream.

6.3 Data Collection Tools

Several instruments were used in Study 2. The main instrument was the pre- and post- writing tests and the 30-minute semi-structured interviews with the student and teacher participants. Other instruments included background questionnaires, classroom observations and course outlines (see Appendix I and Appendix J). In addition, students’ individual assignments were collected (see Table 6.3 and 6.4 for the data
collection schedule) at the end of each semester after students had submitted work to the teacher. Each instrument is discussed in detail below.

6.3.1 Pre- and post-writing tests

A 30-minute writing task was used as the pre- and post-test for Phase 1 and a post-test for Phase 2 of this study. The same topic was used for the pre- and post-test in each phase. Using the same topics in both tests aligns with recommended pre-and post-test research design (e.g., Elder & O’Loughlin, 2003).

The writing task had no word limit. Topics for the tests were based on the writing tasks in Japanese 3. The objective of Japanese 3 was for students to introduce a place in Japanese to a particular audience by the end of semester. The pre- and post-test writing topic required students to write a promotional newspaper/magazine article about their hometown or a place they knew very well (see Appendix K for full details).

6.3.2 Semi-structured interviews

Semi-structured interviews were conducted in English with both the student and teacher participants. The interviews were approximately 30 minutes in length and were conducted in a quiet area in the university library with students and in the university office with the teachers. The interviews were digitally recorded using an Olympus Digital Voice Recorder VN-733PC and were transcribed by the researcher. The average interview length was 24 minutes.

The questions in the interviews elicited students’ perspectives on their writing development and their views about the assignments and writing activities done both inside and outside class. In Japanese 3, the researcher attempted to elicit whether tasks such as the small cultural group projects, weekly kanji tests and weekly reflective writing tasks helped learners improve their writing. General questions were also asked, such as the extent of participants’ enjoyment of writing in their first language/L2 Japanese and contact with native speakers of Japanese. As the classroom activities and
assignments differed slightly between Japanese 3 and 4, the information the researcher tried to obtain in interviews with participants in Japanese 4 varied from Japanese 3. In Japanese 4, apart from questions related to group assignments and *kanji* tests, the researcher also focused on obtaining information related to the peer feedback undertaken for both in-class activities and assignments (see Appendix L for Japanese 3 interview questions and Appendix M for Japanese 4 interview questions).

For both Japanese 3 and Japanese 4, semi-structured interviews were also conducted with the teachers at the end of the respective semesters (approximately 30 minutes’ duration). In Japanese 3, two teachers (one sessional tutor and one lecturer/coordinator) participated in the interview, which was also conducted in English. In Japanese 4, only the lecturer/coordinator took part in the interview. The interview questions were related to assignments, writing activities and their perceptions of students’ writing development, and the impact of language background on writing development. In Japanese 3, the researcher attempted to elicit what aspects of class work and assignments supported students’ writing development. General questions were also included, such as learning materials outside class and students’ motivation towards writing. As with the student interviews, some questions relating to classroom activities and assessments were different in the two interviews. In Japanese 4, apart from responses related to assignments and *kanji* tests, the researcher also tried to obtain information related to the peer feedback assessments (see Appendix N for Japanese 3 interview questions and Appendix O for Japanese 4). In early 2016, a follow-up interview (30 minutes) was also conducted with the coordinator. The questions mainly related to the course design. They also asked whether the coordinator had made any changes to the curriculum in the year following the researcher’s data collection (see Appendix P).
6.3.3 Background questionnaire

Background questionnaires (also used for Study 1) were completed by the participants in Japanese 3. The questions asked for personal details (e.g., name, student ID, faculty), language background (e.g., first language, whether any family member spoke Japanese) and education in Japanese (e.g., if they had studied Japanese in Japan, and if they had completed Year 12 in Australian school/overseas) (see Appendix B).

6.3.4 Lecture and classroom observation

(Week 3–Week 10 in Japanese 3 and Japanese 4)

The researcher attended and observed one lecture and one class every week in both Japanese 3 and Japanese 4. The lectures did not focus on writing. The lectures conveyed information about Japanese culture in Japanese 3 and the Japanese writing system (e.g., how kanji became part of Japanese) in Japanese 4. The classes that the researcher observed weekly (Seminar 2) incorporated writing activities (although they were very limited). The focus of the observation was to see what type of writing was done in class and the nature of writing instruction. In both subjects, the researcher was a participant observer and was therefore able to interact with the learners, including helping them when they had questions. In Japanese 3, the researcher assisted the tutor with giving feedback to students for the 10-minute writing task done at the end of each class. As the focus of the observation was only on writing and there were very limited opportunities for learners to practice their writing in class, observation schema were not used. Only notes relating to the students’ writing practices (e.g., how much writing the students did) and writing instruction were taken.

A typical classroom activity (for both Japanese 3 and Japanese 4) consisted of the teacher going through the textbook material and asking students to answer questions found there, which took about 20 minutes of the class time. Several grammar structures were taught in Japanese 3 and Japanese 4; for example, conjunctive particles such as ga
(but) and node (because), conjugational forms such as reba-form (if), the tari-form (and). These were taught explicitly. At the end of each class, a 10-minute reflective writing task was undertaken, but these were very short and limited to three sentences.

Prior to writing their assignments in Japanese 3 and Japanese 4, teachers also briefly described how to organise an essay. For example, in Week 6, the concept of coherence was taught explicitly in Japanese 3, using power point slides that showed examples of coherence. After this instruction, the students completed exercises using workbooks. Japanese 4 focused on teaching various expressions; honorific, humble and polite expressions were explicitly taught.

The teaching approach observed predominantly incorporated the product approach, with occasional elements of the genre approach. The product approach was characterised by classroom practices such as weekly grammar instruction and workbook exercises (e.g., substitution exercises). The genre approach in Japanese 3 was characterised by teaching learners the features of tourism texts.

6.3.5 Individual assignments

In both Japanese 3 and Japanese 4, students completed individual assignments. Each subject included only one individual assignment; therefore, the purpose of collecting assignments was to examine whether students used the various expressions taught in class during the semester/s. It was not possible to examine developmental changes using the assignments, only one assignment was issued per semester and each assignment had a different topic.

6.4 Data-Coding Procedures

The current study employed discourse analytic measures chosen according to the findings of Study 1, as well as rating scales. Classroom observations and assessment tasks in Study 2 enabled the researcher to explore writing features other than those examined in Study 1; therefore, rating scales were an additional feature in Study 2. The
initial analysis of the writing scripts only used the discourse analytic measures from Study 1. However, this approach was not sufficient to capture the quality of writing produced; therefore, rating scales were developed for two of the additional features, content and structure.

To ensure that a broad range of writing aspects was analysed, the research drew on Knoch’s (2011) taxonomy of writing features. Knoch showed that no model of text production or writing proficiency is by itself sufficient to represent what learners produce in a written text. Thus, Knoch has proposed a taxonomy of features based on the available theories and models: accuracy, fluency, complexity, mechanics, cohesion, coherence, reader/writer interaction and content.

Knoch’s proposed taxonomy features also partially informed the development of the scales used for this study. Out of this categorisation, CAF were used as the discourse analytic measures in this study. The discourse measures from Study 1 that could be objectively coded are shown in Table 6.6. The coding for these measures is described in detail in Chapter 4 (Section 4.4.1.1). The other features, which were more difficult to capture quantitatively, were captured by a series of rating scales designed for this study, to suit the writing task. Two main features (with two constructs each) were considered for the rating scales: content (sufficiency and persuasiveness) and structure (coherence and paragraphing).

Table 6.6

*Discourse Analytic Measures*

<table>
<thead>
<tr>
<th>Construct</th>
<th>Discourse Analytic Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kanji complexity</td>
<td>KC RTTR, KC CTTR, KW RTTR, KW CTTR</td>
</tr>
<tr>
<td>Syntactic complexity</td>
<td>Ratio of te-form, ratio of complex sentences, clause conjunctions</td>
</tr>
<tr>
<td></td>
<td>ratio of propositional endings, ratio of epistemic endings</td>
</tr>
<tr>
<td>Accuracy</td>
<td>Error-free sentences, EFCs, words per error-free sentence</td>
</tr>
<tr>
<td>Fluency</td>
<td>Number of words, sentences, clauses, KCs, KWs</td>
</tr>
</tbody>
</table>
In the next section, a description of the rating scale used in this study will be introduced.

6.4.1 Content scale

In general, content refers to the ideas communicated through a text. Many discourse measures of content in the literature (e.g., Jacobs, Zinkgraf, Wormuth, Hartfiel & Hughey, 1981; Scarino et al., 2011; Storch & Tapper, 2009) are either not very detailed or descriptions in the criteria are ambiguous. For example, in Scarino et al. (2011) the content was concerned only with information (i.e., whether the information was detailed or not). Jacobs et al. (1981) developed a weighted L2 5-descriptor writing scale that included content, with scores ranging from 13 (very poor) to 30 (excellent). Nevertheless, some of the descriptions are ambiguous—terminologies such as ‘substantive’, ‘little substantive’, ‘inadequate development of topic’, ‘thorough development of thesis’ were not distinguished from each other and were not very clear. The researcher therefore developed a detailed 4-level scale with descriptors (see Table 6.7).

Considering that the genre for this study’s writing task was promotional material, two constructs were chosen for content: sufficiency and persuasiveness.

Relevance was initially considered as a measure; however, after an analysis of 10 per cent of the Japanese 3 pre-test writing scripts, it was deemed unnecessary. This was because in most cases, the writing that students produced was related to the topic. Therefore, content was more concerned with the detail of the relevant material that they provided being sufficient.

A brief definition of each construct is presented (with an explanation of how other studies have informed the scoring criteria) followed by a summary of how they were assessed. Examples of the coding are given in Appendix Q.
6.4.1.1 Sufficiency

Sufficiency refers to the amount of information and detail provided by the writer. The rating scales were partially informed by Scarino et al. (2011, p. 685). The scoring criteria for content used for the independent writing task (self-introduction) in Scarino et al. (2011) was concerned with the level of detail in information that students provided. A 5-level descriptor was used for the scoring criteria, ranging from ‘0’ (no response, few words) to ‘4’ (an extensive range of information was included). This study used a 4-level descriptor to assess sufficiency; a score of ‘3’ was given if an extensive range of information was included that related to the topic. A score of ‘2’ was given if a good range of information was provided but the explanation was brief. A score of ‘1’ indicated that a limited range of information was provided and ‘0’ if there was no response or the learner wrote very few words.

6.4.1.2 Persuasiveness

Persuasiveness refers to the way in which information is provided to the reader, to convince the reader to take on the writer’s perspective. The writing task for this study was to produce a promotional newspaper or article about a place that the writer knew well. Persuasiveness was therefore examined by determining if the writer included reasons regarding why a place was worth visiting or whether he or she featured something unique or special about the place. Ideas for this rating scale were partially informed by the various criteria in the Naplan Persuasive Writing Marking Guide (ACARA, 2013) (e.g., audience, text structure and ideas). For example, the criteria for ideas in Naplan were concerned with how well the idea was elaborated. In this study, persuasiveness was examined in relation to how well the writer elaborated appropriate reasons for visiting the chosen place.
<table>
<thead>
<tr>
<th>Construct</th>
<th>Score</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sufficiency</td>
<td></td>
<td>An extensive range of information is included (e.g., the name of the town, what there is to do, how to get to the town attractions, how much it costs to get there). Each piece of information has sufficient detail.</td>
<td>A good range of information is included (e.g., the name of the town, what there is to do, how to get to the town attractions, how much it costs to get there) but each piece of information may be brief and lacking detail.</td>
<td>Limited range of information is included (e.g., the name of the town, what there is to do) and each piece of information may be brief and lacking detail.</td>
<td>No response or just a few words.</td>
</tr>
<tr>
<td>Persuasiveness</td>
<td></td>
<td>Grabs the attention of the reader by giving appropriate reasons/convincing evidence for visiting the place.</td>
<td>Gives an appropriate and adequate number of reasons but the information may be brief/reason not supported appropriately.</td>
<td>Gives appropriate reasons for visiting the place but may be limited to one or two reasons.</td>
<td>Gives no reasons for why the place is worth visiting.</td>
</tr>
</tbody>
</table>
6.4.2 Structure

Structure was divided into two constructs: coherence and paragraphing. Some literature suggests (Knoch, 2009, 2011; Polio, 2001) that paragraphing should be considered part of mechanics. Nevertheless, some scoring rubrics (e.g., IELTS writing scoring rubric) combine coherence/cohesion/paragraphing into one construct. In this study, coherence and paragraphing were separated, as some students were able to write coherently but only wrote one paragraph. Hence, to examine logical development (i.e., coherence) and paragraphing it was necessary to separate the two criteria. Descriptions of the two constructs are discussed below and coding examples are given in Appendix R.

6.4.2.1 Coherence

Coherence refers to the logical development of ideas in the text and how organised the text is. To assess coherence, a score of ‘3’ was given when all information/ideas flowed smoothly, ‘2’ when most sentences flowed smoothly, but in some cases the information between paragraphs or sentences were difficult to follow, ‘1’ when there was a lack of coherence within or/and between paragraphs, and ‘0’ when the entire text was difficult to follow. Table 6.8 shows the scale for coherence/logical development.

6.4.2.2 Paragraphing

Paragraphing refers to the way in which ideas are presented by either indenting or leaving spaces between paragraphs. By organising ideas in an effective manner through paragraphing, the writer shows how well he or she understands the connections between ideas.

The number of bands was determined by examining 10 per cent of the Japanese 3 scripts (pre- and post-test). A clear distinction was evident between those
students who used paragraphs appropriately, those who made some effort, but sometimes had problems grouping ideas together into paragraphs, and those who wrote only one long paragraph. Table 6.9 shows the scale for paragraphing.

Table 6.8

*Coherence/Logical Development*

<table>
<thead>
<tr>
<th>Score</th>
<th>Construct</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coherence/logical development</td>
</tr>
<tr>
<td>3</td>
<td>Main point is clear, information is organised logically; ideas are smoothly linked throughout the text.</td>
</tr>
<tr>
<td>2</td>
<td>Most ideas are located where they should be. Most information is organised logically but sometimes information is difficult to follow.</td>
</tr>
<tr>
<td>1</td>
<td>Several ideas are not located where they should be. Presents information with limited organisation &amp; flow; several ideas not linked smoothly to each other.</td>
</tr>
<tr>
<td>0</td>
<td>Entire text difficult to follow.</td>
</tr>
</tbody>
</table>

Table 6.9

*Paragraphing*

<table>
<thead>
<tr>
<th>Score</th>
<th>Construct</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Paragraphing</td>
</tr>
<tr>
<td>2</td>
<td>There are clearly distinguished paragraphs and related ideas are appropriately grouped into paragraphs.</td>
</tr>
<tr>
<td>1</td>
<td>Some attempt at paragraphing. Some problems grouping ideas into paragraphs.</td>
</tr>
<tr>
<td>0</td>
<td>No paragraphing at all.</td>
</tr>
</tbody>
</table>

6.5 *Reliability Analysis*

Inter-rater reliability for the discourse analytic measures was not necessary in Study 2, as it had already been undertaken in Study 1 (see Chapter 4, Section 4.4.4). However, inter-rater reliability was necessary for the rating scales. Ten per cent of
the writing scripts (pre- and post-tests) from Japanese 3 and Japanese 4 (11 out of 103) were randomly selected for inter-rater reliability. The second coder, a PhD student in applied linguistics who is a native Japanese speaker, coded the scripts using the two rating scales of content and structure. To ensure understanding of concepts, such as sufficiency and persuasiveness, the second coder was trained by coding three scripts, reviewing the coding guidelines, and discussing the coding procedure with the researcher. The Spearman’s Rho statistics for both content (sufficiency and persuasiveness) and structure (coherence and paragraphing) were above .90. The percentage of agreement for content (both sufficiency and persuasiveness) was 92 per cent and the agreement for structure was 90 per cent for coherence and 92 per cent for paragraphing. This indicates that the measures used for rating scales were all reliable.

6.6 Data Analysis Procedure

The following section presents information on how the discourse analytic measures, rating scales and interviews were analysed.

6.6.1 Discourse analytic measures and rating scales analysis

The pre- and post-test data of the 41 participants in Japanese 3 were used to investigate development over one semester (short term). Twenty-two participants who continued to study Japanese 4 undertook a post-test writing test. Data from these 22 participants (from Japanese 3 pre-test and Japanese 4 post-test) were used to examine development over one year (long term). As mentioned before, one of the study’s aims was to establish if an improvement in writing over time existed, and if so, which areas developed. The features considered to measure development in Study 2 were kanji complexity, syntactic complexity, accuracy, fluency, content and structure. In addition, data from the Japanese 3 group were divided into character-
based learners and non-character-based learners in order to investigate the differences (if any) in development between the two groups.

The analysis for the pre- and post-test writing tasks was undertaken for the results after one semester and after one year. First, descriptive statistics were obtained using SPSS. Repeated measures ANOVA were performed for the pre- and post-test writing tasks to determine if the discourse analytical measures (see Table 6.6) were statistically significant. Assumptions of normality, homogeneity of variance, linearity, homogeneity, and reliability of covariates were tested for each measure; these were satisfactory.

A similar procedure was used to analyse the data for character-based learners and non-character-based learners. First, descriptive statistics were obtained. Then, repeated measures ANOVA were performed. When the results showed significant differences between the two groups, pairwise comparison post-hoc tests were conducted to determine the difference in the rate of improvement between the two groups.

Effect sizes were also considered when reporting the results. Based on Cohen (1998), the $\eta^2$ (Eta squared) effect sizes were interpreted as follows:

- .01 = small
- .06 = moderate
- .14 = large.

6.6.2 Analysing interviews

Semi-structured interviews were conducted with 40 students and two teachers (lecturer and tutor) in Japanese 3, and 22 students and the lecturer in Japanese 4. The main aim of the interviews was to gain insights about the participants’ perspectives.
and ask them what kind of activities and assessment tasks helped to improve their writing.

Interviews conducted in Japanese 3 and Japanese 4 were transcribed by the researcher and saved as text files. Student interviews and teacher interviews were analysed separately. Analysis of the student interviews is discussed here, followed by an analysis of the teacher interviews.

6.6.2.1 Analysis of student interviews

After transcribing the interviews and thoroughly reading and re-reading the data, themes and sub-themes were identified. This yielded three broad themes (each consisting of several sub-themes). The broad themes (and sub-themes) identified from student interviews are shown in Table 6.10. Each theme was analysed and discussed under a few sub-themes with examples (i.e., quotations) from the interviews. However, some themes were only relevant to Japanese 3, while others were relevant to both Japanese 3 and Japanese 4. The themes are discussed accordingly.

Table 6.10

Broad Themes and Sub-themes Identified for Student Interviews

<table>
<thead>
<tr>
<th>Themes</th>
<th>Sub-themes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitudes to writing in L1 and L2</td>
<td>Reasons why students like/dislike writing in L1</td>
</tr>
<tr>
<td></td>
<td>Reasons why students like/dislike writing in L2 (Japanese)</td>
</tr>
<tr>
<td>Self-assessment of writing improvement</td>
<td>Grammar improvement</td>
</tr>
<tr>
<td></td>
<td>Vocabulary</td>
</tr>
<tr>
<td></td>
<td>Kanji improvement</td>
</tr>
<tr>
<td>Factors which influenced students’</td>
<td>Practice in writing</td>
</tr>
<tr>
<td>writing development</td>
<td>Feedback</td>
</tr>
<tr>
<td></td>
<td>Teacher feedback</td>
</tr>
<tr>
<td>Themes</td>
<td>Sub-themes</td>
</tr>
<tr>
<td>--------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Peer feedback:</td>
<td>Non-assessed peer feedback (group work)</td>
</tr>
<tr>
<td></td>
<td>Assessed peer feedback</td>
</tr>
<tr>
<td>Reading in Japanese</td>
<td></td>
</tr>
</tbody>
</table>

### 6.6.2.2 Analysis of teacher interviews

Similar to the student interview analysis, the transcribed interviews were read thoroughly and themes and sub-themes identified. The responses were then classified into four broad themes (see Table 6.11). Each theme was analysed and discussed under a few sub-themes with examples (i.e., quotations) from the interviews.

Table 6.11

*Broad Themes and Sub-themes Identified for Teacher Interviews*

<table>
<thead>
<tr>
<th>Themes</th>
<th>Sub-themes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overview of the course design</td>
<td></td>
</tr>
<tr>
<td>Teachers’ views on L2 writing and writing development</td>
<td>Importance of teaching of Japanese writing</td>
</tr>
<tr>
<td>Factors that influenced students’ progress in writing</td>
<td>Practice in writing</td>
</tr>
<tr>
<td></td>
<td>Reading in Japanese</td>
</tr>
<tr>
<td></td>
<td>Feedback on writing</td>
</tr>
<tr>
<td>Student background</td>
<td>Character-based learners and non-character-based learners</td>
</tr>
</tbody>
</table>

The following chapter presents the short- and long-term results in students’ writing development, as well as the results of the differences between character-based learners and non-character-based learners.
Chapter 7: Study 2 Text-Based Analyses Results

The purpose of Study 2 was to examine the writing development of L2 Japanese lower-intermediate learners in the short term (over one semester) and in the long term (over one year) (Study 2, RQ 1). Another aim was to determine whether a difference in writing development existed between the character-based learners and non-character-based learners (Study 2, RQ 2). Thus, various writing constructs were used to examine whether writing improved over time: *kanji* complexity, syntactic complexity, accuracy, fluency, content and structure.

Study 2 comprised 41 participants in Japanese 3 and 22 in Japanese 4. All 41 students first wrote a pre-test at the beginning of Japanese 3 and wrote again at the end of Japanese 3. Twenty-two of these students continued their studies in Japanese 4 and completed another piece of writing (on the same topic as Japanese 3) at the end of the Japanese 4. Short-term results are those from Japanese 3 pre- and post-tests. Long-term results are those from the Japanese 3 pre-test (only using students who continued to Japanese 4; i.e., \( N = 22 \)) and the Japanese 4 post-test.

This chapter presents the descriptive and inferential results for each construct investigated for Study 2. Therefore, in accordance with RQ1 for Study 2, results for the short- and long term are provided. This is followed by the results for character-based learners and non-character-based learners, which reflects RQ2 for Study 2.

7.1 *Kanji* Complexity: Short- and Long-term Results

To determine whether an improvement in *kanji* complexity existed over the short- and long term, various measures were investigated: KC RTTR, KC CTTR, KW RTTR and KW CTTR. Table 7.1 displays the descriptive statistics for the short- and long term for *kanji* complexity.
All measures examined for *kanji* complexity showed an increase from pre- to post-test in both the short and long term (see Table 7.1). For example, in the short term, KC RTTR increased by 1.18 and in the long term it showed an increase of 1.86. Moreover, the increase in the mean for each measure in the long term was greater compared to that of the short term. Hence, the results suggest that students’ Japanese L2 writing developed in terms of *kanji* complexity in the long term more than in the short term.

Table 7.1

*Descriptive Statistics for the Short and Long Term: Kanji Complexity*

<table>
<thead>
<tr>
<th>Measures</th>
<th>Short term</th>
<th>Long term</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test</td>
<td>Post-test</td>
</tr>
<tr>
<td></td>
<td>(N = 41)</td>
<td>(N = 41)</td>
</tr>
<tr>
<td>KC RTTR</td>
<td>3.67</td>
<td>4.85</td>
</tr>
<tr>
<td></td>
<td>1.12</td>
<td>1.26</td>
</tr>
<tr>
<td>KC CTTR</td>
<td>2.62</td>
<td>3.40</td>
</tr>
<tr>
<td></td>
<td>0.79</td>
<td>0.92</td>
</tr>
<tr>
<td>KW RTTR</td>
<td>3.37</td>
<td>4.42</td>
</tr>
<tr>
<td></td>
<td>0.88</td>
<td>0.90</td>
</tr>
<tr>
<td>KW CTTR</td>
<td>2.38</td>
<td>3.13</td>
</tr>
<tr>
<td></td>
<td>0.63</td>
<td>0.63</td>
</tr>
</tbody>
</table>

To establish whether any statistically differences existed between the pre- and post-tests for the short and long term, repeated measure ANOVAs were performed. The results are shown in Table 7.2.

Table 7.2

*Repeated Measures ANOVA Results for the Short and Long Term: Kanji Complexity*

<table>
<thead>
<tr>
<th>Measures</th>
<th>Short-term results</th>
<th>Long-term results</th>
</tr>
</thead>
<tbody>
<tr>
<td>KC RTTR</td>
<td>F(1, 40) = 47.81, p &lt; .001, $\eta^2 = .54$</td>
<td>F(1, 21) = 48.45, p &lt; .001, $\eta^2 = .70$</td>
</tr>
<tr>
<td>KC CTTR</td>
<td>F(1, 40) = 37.01, p &lt; .001, $\eta^2 = .48$</td>
<td>F(1, 21) = 44.59, p &lt; .001, $\eta^2 = .68$</td>
</tr>
<tr>
<td>KW RTTR</td>
<td>F(1, 40) = 50.59, p &lt; .001, $\eta^2 = .56$</td>
<td>F(1, 21) = 26.99, p &lt; .001, $\eta^2 = .56$</td>
</tr>
<tr>
<td>KW CTTR</td>
<td>F(1, 40) = 51.88, p &lt; .001, $\eta^2 = .57$</td>
<td>F(1, 21) = 28.07, p &lt; .001, $\eta^2 = .57$</td>
</tr>
</tbody>
</table>
Table 7.2 indicates that both KC RTTR and KC CTTR showed significant differences between the pre- and post-test (p < .001) in both the short and long term, with a large effect size. Similarly, KW RTTR and KW CTTR showed statistically significant difference between pre-test and post-test (p < .001) in both the short and long term with a large effect size. The results suggest that by spending more time learning Japanese, students can use a wider variety of KCs and KWs in their writing.

Although the effect size for all measures had a large effect size, the effect size for KC RTTR and CTTR in the long term was slightly higher compared to the other measures. This means that the improvement for KC RTTR and CTTR in the long term was slightly better compared to other measures; for example, KW RTTR in the short and long term.

7.2 Syntactic Complexity: Short- and Long-term Results

To examine the writing development of syntactic complexity, six different measures were considered: ratio of te-form, ratio of complex sentences, ratio of clause conjunctions, ratio of propositional endings, ratio of epistemic endings and ratio of expressing endings. Short- and long-term results for each of these measures are given in Table 7.2.3. As this shows, very little change was evident for syntactic complexity in most of the measures in both the short and long term. Noticeable improvements were made in relation to ratio of complex sentences and ratio of clause conjunctions.
Table 7.3

**Descriptive Statistics for Short and Long Term: Syntactic Complexity**

<table>
<thead>
<tr>
<th>Measures</th>
<th>Short term</th>
<th>Long term</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test (N = 41)</td>
<td>Post-test (N = 41)</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Ratio of te-form</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Ratio of complex sentences</td>
<td>0.14</td>
<td>0.16</td>
</tr>
<tr>
<td>Ratio of clause conjunction</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>Ratio of propositional endings</td>
<td>0.07</td>
<td>0.03</td>
</tr>
<tr>
<td>Ratio of epistemic endings</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>Ratio of expressive endings</td>
<td>0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>

As seen in Table 7.4, in the short term, the measures with statistically significant differences were the ratio of te-form, the ratio of complex sentences and the ratio of clause conjunctions. Te-form had a medium effect size, while complex sentences and ratio of clause conjunctions had large effect sizes. In the long term, only the ratio of complex sentences and clause conjunctions showed significant differences between the pre- and post-test, with a large effect size. Expressive endings also showed significant differences (p = .045) with a large effect size ($\eta^2 = .18$). However, with this measure the mean number had decreased, which indicates that students’ performance in this area deteriorated after a year. The difference in mean number was only 0.01; therefore, even though there was a significant difference, a definitive conclusion cannot be made. The above results suggest that one year of language study may not be enough for students to improve in syntactic complexity.
Table 7.4

Repeted Measure ANOVA Results for Short and Long Term: Syntactic Complexity

<table>
<thead>
<tr>
<th>Measures</th>
<th>Short-term results</th>
<th>Long-term results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Te-form</td>
<td>$F(1, 40) = 4.84, p = .030, \eta^2 = .11$</td>
<td>$F(1, 21) = 0.06, p = .806, \eta^2 = \text{n/a}$</td>
</tr>
<tr>
<td>Complex sentences</td>
<td>$F(1, 40) = 8.50, p = .010, \eta^2 = .18$</td>
<td>$F(1, 21) = 4.76, p = .041, \eta^2 = .19$</td>
</tr>
<tr>
<td>Clause conjunction</td>
<td>$F(1, 40) = 7.28, p = .010, \eta^2 = .15$</td>
<td>$F(1, 21) = 23.40, p &lt; .001, \eta^2 = .53$</td>
</tr>
<tr>
<td>Propositional endings</td>
<td>$F(1, 40) = 3.41, p = .070, \eta^2 = \text{n/a}$</td>
<td>$F(1, 21) = 3.04, p = .096, \eta^2 = \text{n/a}$</td>
</tr>
<tr>
<td>Epistemic endings</td>
<td>$F(1, 40) = .109, p = .740, \eta^2 = \text{n/a}$</td>
<td>$F(1, 21) = 0.41, p = .531, \eta^2 = \text{n/a}$</td>
</tr>
<tr>
<td>Expressive endings</td>
<td>$F(1, 40) = .325, p = .520, \eta^2 = \text{n/a}$</td>
<td>$F(1, 21) = 4.57, p = .045, \eta^2 = .18$</td>
</tr>
</tbody>
</table>

7.3 Accuracy: Short- and Long-term Results

Writing development in accuracy was examined by investigating three different measures: the ratio of error-free sentences, the ratio of EFCs and words per error-free sentence. Descriptive statistics for these measures are displayed in Table 7.5. As seen in this table, the mean number for all measures of accuracy increased from pre- to post-test in the short term, whereas the mean number for all measures of accuracy declined from pre- to post-test in the long term.

Table 7.5

Descriptive Statistics for Short and Long Term: Accuracy

<table>
<thead>
<tr>
<th>Measures</th>
<th>Short term Pre-test ($N = 41$)</th>
<th>Short term Post-test ($N = 41$)</th>
<th>Long term Pre-test ($N = 22$)</th>
<th>Long term Post-test ($N = 22$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>SD</td>
<td>$M$</td>
<td>SD</td>
</tr>
<tr>
<td>Ratio of error-free sentences</td>
<td>0.30</td>
<td>0.21</td>
<td>0.36</td>
<td>0.20</td>
</tr>
<tr>
<td>Ratio of EFCs</td>
<td>0.32</td>
<td>0.28</td>
<td>0.46</td>
<td>0.77</td>
</tr>
<tr>
<td>Words per error-free sentence</td>
<td>6.85</td>
<td>3.65</td>
<td>8.98</td>
<td>4.61</td>
</tr>
</tbody>
</table>
Table 7.6 shows the results from repeated measures ANOVA for the short and long term. In the short term, the measures with significant differences between pre- and post-test were the ratio of EFCs and words per error-free sentence, with a medium effect size for the ratio of EFCs and a large effect size for words per error-free sentence. In the long term, the only measure that showed significant differences between pre- and post-test was the ratio of error-free sentences, with a large effect size. Nevertheless, the mean number decreased from pre- to post-test (see Table 7.5), indicating that instead of improving, students’ performance had worsened on this particular measure. The results suggest that accuracy may fluctuate: it can improve in the short term but deteriorate in the long term. Possible reasons for this are discussed in Section 9.1.3 (in relation to Study 2).

Table 7.6

<table>
<thead>
<tr>
<th>Measures</th>
<th>Short-term results</th>
<th>Long-term results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratio of error-free sentences</td>
<td>F(1, 40) = 6.15, p = .282, η² = n/a</td>
<td>F(1, 21) = 5.03, p = .036, η² = .19</td>
</tr>
<tr>
<td>Ratio of EFCs</td>
<td>F(1, 40) = 1.19, p = .017, η² = .030</td>
<td>F(1, 21) = 0.98, p = .334, η² = n/a</td>
</tr>
<tr>
<td>Words per error-free sentence</td>
<td>F(1, 40) = 7.37, p = .010, η² = .16</td>
<td>F(1, 21) = 0.01, p = .925, η² = n/a</td>
</tr>
</tbody>
</table>

7.4 Fluency: Short- and Long-term Results

Five different measures were considered to determine writing development in fluency for the short and long term: the number of words, the number of sentences, the number of clauses, the number of KCs and the number of KWs. Table 7.7 illustrates the descriptive statistics for fluency.
Table 7.7 shows that an improvement existed in all the fluency measures in both the short and long term; the mean number for all fluency measures increased from pre- to post-test. The difference in the mean number for pre- and post-test in the short term was greater for the number of words, sentences and clauses. For example, the mean for the number of words on the pre-test was 92.66 and 156.49 on the post-test in the short term, indicating an increase of 63.83. However, the mean for the number of words in the pre-test was 90.96 and 147.77 in the post-test in the long term, showing a 56.81 increase.

Table 7.7

*Descriptive Statistics for the Short and Long Term: Fluency*

<table>
<thead>
<tr>
<th>Measures</th>
<th>Short term</th>
<th></th>
<th>Long term</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test (N = 41)</td>
<td>Post-test (N = 41)</td>
<td>Pre-test (N = 22)</td>
<td>Post-test (N = 22)</td>
</tr>
<tr>
<td>No. of words</td>
<td>92.66</td>
<td>33.87</td>
<td>156.49</td>
<td>51.06</td>
</tr>
<tr>
<td>No. of sentences</td>
<td>9.66</td>
<td>3.43</td>
<td>14.39</td>
<td>4.74</td>
</tr>
<tr>
<td>No. of clauses</td>
<td>5.24</td>
<td>3.14</td>
<td>9.49</td>
<td>4.91</td>
</tr>
<tr>
<td>No. of KCs</td>
<td>27.22</td>
<td>16.50</td>
<td>57.68</td>
<td>32.23</td>
</tr>
<tr>
<td>No. of KWs</td>
<td>18.54</td>
<td>9.38</td>
<td>35.85</td>
<td>16.11</td>
</tr>
</tbody>
</table>

The repeated measures ANOVA results in Table 7.8 reveal that all measures for fluency show statistically significant differences between the pre- and post-test in both the short and long term, with a large effect size. In the long term, although all measures showed a large effect size, the number of words, KCs and KWs had larger effect sizes than the remaining two measures (i.e., the number of sentences and clauses). This means that the improvement for the number of words, KCs and KWs was greater compared to that for the number of sentences and clauses in the long term.
Table 7.8

Repeated Measures ANOVA Results for Short and Long Term: Fluency

<table>
<thead>
<tr>
<th>Measures</th>
<th>Short-term results</th>
<th>Long-term results</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of words</td>
<td>$F(1, 40) = 74.38, p &lt; .001, \eta^2 = .65$</td>
<td>$F(1, 21) = 29.54, p &lt; .001, \eta^2 = .58$</td>
</tr>
<tr>
<td>No. of sentences</td>
<td>$F(1, 40) = 41.16, p &lt; .001, \eta^2 = .51$</td>
<td>$F(1, 21) = 6.51, p = .019, \eta^2 = .24$</td>
</tr>
<tr>
<td>No. of clauses</td>
<td>$F(1, 40) = 35.95, p &lt; .001, \eta^2 = .47$</td>
<td>$F(1, 21) = 10.95, p = .003, \eta^2 = .34$</td>
</tr>
<tr>
<td>No. of KCs</td>
<td>$F(1, 40) = 52.29, p &lt; .001, \eta^2 = .57$</td>
<td>$F(1, 21) = 21.69, p &lt; .001, \eta^2 = .51$</td>
</tr>
<tr>
<td>No. of KWs</td>
<td>$F(1, 40) = 58.89, p &lt; .001, \eta^2 = .60$</td>
<td>$F(1, 21) = 39.13, p &lt; .001, \eta^2 = .65$</td>
</tr>
</tbody>
</table>

7.5 Content: Short- and Long-term Results

Content was one of the features that could not be coded objectively using discourse analytic measures. Accordingly, rating scales were used to examine the students’ writing development. As described in the previous chapter, two measures were considered for content: sufficiency and persuasiveness. Scores from ‘0’ to ‘3’ were assigned to each of the two measures.

Table 7.9 shows the descriptive statistics for mean scores and standard deviations in the short and long term for content. As shown in Table 7.9, students showed improvement in sufficiency and persuasiveness in both the short and long term. Nevertheless, the increase in the mean scores over the long term was slightly smaller (0.95 for sufficiency and 0.77 for persuasiveness) compared to that for the short-term results (0.98 for sufficiency and 0.81 for persuasiveness).
Table 7.9

Descriptive Statistics for Short and Long Term: Content ( Sufficiency and Persuasiveness )

<table>
<thead>
<tr>
<th>Measures</th>
<th>Short term</th>
<th>Long term</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test (N = 41)</td>
<td>Post-test (N = 41)</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Sufficiency</td>
<td>1.22</td>
<td>0.61</td>
</tr>
<tr>
<td>Persuasiveness</td>
<td>1.17</td>
<td>0.50</td>
</tr>
</tbody>
</table>

To determine whether the changes from the pre- to post-test were statistically significant, repeated measures ANOVAs were performed. The results reveal that a statistically significant difference existed for changes in the short and long term, with a large effect size for both sufficiency and persuasiveness (see Table 7.10).

Table 7.10

Repeated Measures ANOVA Results for the Short and Long Term: Content

<table>
<thead>
<tr>
<th>Measures</th>
<th>Short-term results</th>
<th>Long-term results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F(1, 40) = 67.94, p &lt; .001, ( \eta^2 = .63 )</td>
<td>F(1, 21) = 32.50, p &lt; .001, ( \eta^2 = .61 )</td>
</tr>
<tr>
<td>Sufficiency</td>
<td>( \eta^2 = .59 )</td>
<td>( \eta^2 = .57 )</td>
</tr>
<tr>
<td>Persuasiveness</td>
<td>F(1, 40) = 0.41, p &lt; .001,</td>
<td>F(1, 21) = 27.97, p &lt; .001,</td>
</tr>
</tbody>
</table>

7.6 Structure: Short- and Long-term Results

Structure was another feature that could not be coded objectively using discourse analytic measures. As such, rating scales were employed. As mentioned previously, structure was assessed based on a score from ‘0’ to ‘3’ for coherence and ‘0’ to ‘2’ for paragraphing.

Table 7.11 shows the descriptive statistics for mean scores and standard deviations for the short and long term. Learners showed improvement in coherence and paragraphing in both the short and long term. Further, the rate of improvement
was larger for the long term compared to the short term for both coherence and paragraphing.

Table 7.11

*Descriptive Statistics for the Short and Long Term: Structure (Coherence and Paragraphing)*

<table>
<thead>
<tr>
<th>Measures</th>
<th>Short term</th>
<th></th>
<th>Long term</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test</td>
<td>Post-test</td>
<td>Pre-test</td>
<td>Post-test</td>
</tr>
<tr>
<td></td>
<td>(N = 41)</td>
<td>(N = 41)</td>
<td>(N = 22)</td>
<td>(N = 22)</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Coherence</td>
<td>1.12</td>
<td>0.78</td>
<td>1.88</td>
<td>0.78</td>
</tr>
<tr>
<td>Paragraphing</td>
<td>0.59</td>
<td>0.74</td>
<td>1.27</td>
<td>0.71</td>
</tr>
</tbody>
</table>

To determine if the changes were statistically significant, repeated measures ANOVAs were performed. The results reveal that a statistically significant difference existed for changes in the short and long term for both coherence and paragraphing (see Table 7.12). Structure improved in the short term, as well as in the long term for both coherence and paragraphing, with a large effect size. The mean values for both measures were greater for the long-term results, with a larger effect size. This suggests that learners were able to write more coherently and organise their paragraphs more effectively after one year of study (long term) compared to after one semester (short term).

Table 7.12

*Repeated Measures ANOVA Results for the Short and Long Term: Structure*

<table>
<thead>
<tr>
<th>Measures</th>
<th>Short-term results</th>
<th>Long-term results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coherence</td>
<td>F(1, 40) = 34.02, p &lt; .001, $\eta^2 = .46$</td>
<td>F(1, 21) = 59.27, p &lt; .001, $\eta^2 = .71$</td>
</tr>
<tr>
<td>Paragraphing</td>
<td>F(1, 40) = 36.64, p &lt; .001, $\eta^2 = .48$</td>
<td>F(1, 21) = 33.77, p &lt; .001, $\eta^2 = .62$</td>
</tr>
</tbody>
</table>
7.7 Summary of the Main Findings Related to Research Question 1

All the measures investigated for *kanji* complexity improved in both the short and long term. This means that participants were able to use a large variety of KCs and KWs in their writing after one semester, as well as after one year. Improvement was greater after one year compared to the results after one semester, which suggests that more time spent learning Japanese enabled students to write a wider variety of KCs and KWs.

The results for syntactic complexity indicate that not many measures showed improvement in both the short and long term. In the short term, the ratio of te-form, the ratio of clause conjunctions and the ratio of complex sentences showed improvement. In the long term, only the ratio of complex sentences and the ratio of clause conjunctions improved, and there were no improvements in the other measures investigated. This suggests that it may take more than one year for students to improve various aspects of syntactic complexity.

The findings for accuracy show that the ratio of EFCs and words per error-free sentences improved in the short term. The long-term results show that students did not improve their accuracy after one year. These results indicate that accuracy may fluctuate: it can improve in one semester but deteriorate in the next semester. Factors such as classroom instruction and assessment tasks may have affected the short- and long-term results. Student interview results are discussed in the following chapter. They may provide some explanation for this outcome.

The results for fluency indicate that all measures improved in the short, as well as in the long term. Only a slight difference in improvement was apparent between the short- and long-term results. This may be due to students having only 30 minutes in which to complete the writing task—it would have been difficult to
increase the number of words in a script. Nevertheless, it can be said that with more time spent learning Japanese, it is likely that students’ fluency will improve.

The findings related to content show that sufficiency improved over the short and long term. However, the rate of improvement was slightly higher in the short compared to the long term. The results for structure indicate that learners showed improvement in both the short and long term. In contrast to the findings for content, the long-term results here showed a greater improvement compared to the short-term results. This suggests that the more time spent on this area, the greater the improvement for structure.

In conclusion, it can be said that kanji complexity, fluency and structure improves after one semester and even more after one year. Content improved but further improvement was not seen in the long term. Two of the accuracy measures improved only in the short and not in the long term. Moreover, improvement was not seen in few of the measures (i.e., accuracy and syntactic complexity) even after studying for one year. These quantitative results can only show us whether students improve in a particular measure over time. It is important to uncover the possible reasons for students improving in some areas but not others. Thus, the main aim of conducting student and teacher interviews was to reveal some of the reasons behind their development. Interview results are provided in the next chapter and a discussion of the quantitative and interview results is provided in Chapter 9.

7.8 Character-based Learners and Non-character-based Learners

The second research question that Study 2 sought to investigate was if a difference existed in L2 Japanese writing development between character-based learners and non-character-based learners. Out of the 41 participants in Japanese 3, 16 students were character-based learners and 25 students were non-character-based
learners. The number of participants was too small in Japanese 4 \((N = 22)\); therefore, to answer RQ2, only data from Japanese 3 were used.

For each of the measures investigated (kanji complexity, syntactic complexity, accuracy, fluency, content and structure) the development of character-based learners and non-character-based learners was compared. This is discussed below.

**7.8.1 Kanji complexity**

Table 7.13 displays the descriptive statistics for each measure of kanji complexity for character-based learners and non-character-based learners. Figure 7.1 reflects the descriptive results related to kanji complexity. As seen in Table 7.13 and Figure 7.1, both groups showed improvement in all measures for kanji complexity; however, character-based learners showed a higher level in the use of KCs and KWs in the pre-test than did non-character-based learners. Moreover, character-based learners improved more than did the non-character-based learners. For example, there was a 1.68 increase in KC RTTR for character-based learners from the pre- to post-test; however, there was only a 0.87 increase for non-character-based learners (see Table 7.13 and Figure 7.1).
Table 7.13

Descriptive Statistics for Character-based and Non-character-based Learners: Kanji Complexity

<table>
<thead>
<tr>
<th>Measures</th>
<th>Character-based ((n = 16))</th>
<th>Non-character-based ((n = 25))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test</td>
<td>Post-test</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>KC RTTR</td>
<td>4.25</td>
<td>1.01</td>
</tr>
<tr>
<td>KC CTTR</td>
<td>3.00</td>
<td>0.72</td>
</tr>
<tr>
<td>KW RTTR</td>
<td>3.62</td>
<td>0.92</td>
</tr>
<tr>
<td>KW CTTR</td>
<td>2.57</td>
<td>0.64</td>
</tr>
</tbody>
</table>

C = character-based learner; NC = non-character-based learner

Figure 7.1. Estimated marginal means for kanji complexity.

Table 7.14 below shows the results from the repeated measures ANOVA. The results indicate that all measures for kanji complexity show a significant difference in the change over time between the two groups. KC RTTR and KW
CTTR had a medium effect size, while KC CTTR and KW RTTR had a large effect size. A post-hoc pairwise comparison test was conducted to determine the differences in development between the character-based learners and non-character-based learners (see Table 7.15). The test revealed that the character-based learners improved more than did the non-character-based learners. These results suggest that although both groups improved from the pre- to post-test (i.e., both groups were able to write using a larger variety of KCs after one semester), the rate of improvement for character-based learners was greater than that for non-character-based learners.

Table 7.14
Repeated Measures ANOVA Results for Character-based and Non-character-based Learners: Kanji Complexity

<table>
<thead>
<tr>
<th>Measures</th>
<th>Character-based and non-character-based learners</th>
</tr>
</thead>
<tbody>
<tr>
<td>KC RTTR</td>
<td>$F(1, 39) = 6.23, \ p = .017, \ \eta^2 = .13$</td>
</tr>
<tr>
<td>KC CTTR</td>
<td>$F(1, 39) = 7.40, \ p = .010, \ \eta^2 = .16$</td>
</tr>
<tr>
<td>KW RTTR</td>
<td>$F(1, 39) = 6.34, \ p = .016, \ \eta^2 = .14$</td>
</tr>
<tr>
<td>KW CTTR</td>
<td>$F(1, 39) = 5.96, \ p = .019, \ \eta^2 = .13$</td>
</tr>
</tbody>
</table>

Table 7.15
Pairwise Comparison: Mean difference for Kanji Complexity

<table>
<thead>
<tr>
<th>Measures</th>
<th>Character-based ($n = 16$)</th>
<th>Non-character-based ($n = 25$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KC RTTR</td>
<td>1.69 ($p &lt; .001$)</td>
<td>.86 ($p &lt; .001$)</td>
</tr>
<tr>
<td>KC CTTR</td>
<td>1.20 ($p &lt; .001$)</td>
<td>.53 ($p &lt; .001$)</td>
</tr>
<tr>
<td>KW RTTR</td>
<td>1.50 ($p &lt; .001$)</td>
<td>.78 ($p &lt; .001$)</td>
</tr>
<tr>
<td>KW CTTR</td>
<td>1.04 ($p &lt; .001$)</td>
<td>.55 ($p &lt; .001$)</td>
</tr>
</tbody>
</table>

These results suggest that learners’ background (i.e., whether they are character-based learners or non-character-based learners) affected the development
of kanji complexity. A difference was already apparent between the pre-test scores of the two groups—character-based learners had higher scores than non-character-based learners. This difference, together with a further improvement by the character-based learners over the semester, resulted in a further gap between the two groups for this particular measure.

### 7.8.2 Syntactic complexity

The descriptive statistics in Table 7.16 and Figure 7.2 show that the ratio of te-form, the ratio of clause conjunctions and the ratio of complex sentences increased for both character-based learners and non-character-based learners between the pre- and post-test, while the ratio of epistemic endings remained the same. The ratio of propositional endings and expressive endings declined for character-based learners (for propositional endings, a decline in ratio indicates development) but there was no change seen for non-character-based learners.

#### Table 7.16

Descriptive Statistics for Character-based and Non-character-based Learners:

<table>
<thead>
<tr>
<th>Syntactic Complexity</th>
<th>Character-based</th>
<th></th>
<th>Non-character-based</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test</td>
<td>Post-test</td>
<td>Pre-test</td>
<td>Post-test</td>
</tr>
<tr>
<td>----------------------</td>
<td>----------</td>
<td>-----------</td>
<td>----------</td>
<td>-----------</td>
</tr>
<tr>
<td>Measures</td>
<td>(n = 16)</td>
<td></td>
<td>(n = 25)</td>
<td></td>
</tr>
<tr>
<td>Ratio of te-form</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>0.02</td>
</tr>
<tr>
<td>Ratio of complex sentences</td>
<td>0.12</td>
<td>0.21</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>Ratio of clause conjunction</td>
<td>0.03</td>
<td>0.02</td>
<td>0.04</td>
<td>0.02</td>
</tr>
<tr>
<td>Ratio of propositional endings</td>
<td>0.08</td>
<td>0.04</td>
<td>0.06</td>
<td>0.03</td>
</tr>
<tr>
<td>Ratio of epistemic endings</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Ratio of expressive endings</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
<td>0.01</td>
</tr>
</tbody>
</table>
a) Ratio of te-form  

b) Ratio of complex sentences

c) Ratio of clause conjunction  
d) Ratio of propositional endings

e) Ratio of epistemic endings  
f) Ratio of expressive endings

C = character-based learner; NC = non-character-based learner

*Figure 7.2. Estimated marginal means for syntactic complexity.*

As seen in Table 7.17, no statistically significant differences existed in the change over time between the two groups; hence, a post-hoc pairwise comparison test was not necessary.
The results suggest that there was no difference in development between the two groups and thus background had no impact on syntactic complexity.

Table 7.17

*Repeated Measures ANOVA Results for Character-based and Non-character-based Learners: Syntactic Complexity*

<table>
<thead>
<tr>
<th>Measures</th>
<th>Character-based and Non-character-based learners</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratio of te-form</td>
<td>$F(1, 39) = .012, p = .912, \eta^2 = \text{n/a}$</td>
</tr>
<tr>
<td>Ratio of complex sentences</td>
<td>$F(1, 39) = .044, p = .835, \eta^2 = \text{n/a}$</td>
</tr>
<tr>
<td>Ratio of clause conjunction</td>
<td>$F(1, 39) = .153, p = .697, \eta^2 = \text{n/a}$</td>
</tr>
<tr>
<td>Ratio of propositional endings</td>
<td>$F(1, 39) = 6.68, p = .081, \eta^2 = \text{n/a}$</td>
</tr>
<tr>
<td>Ratio of epistemic endings</td>
<td>$F(1, 39) = .002, p = .978, \eta^2 = \text{n/a}$</td>
</tr>
<tr>
<td>Ratio of expressive endings</td>
<td>$F(1, 39) = .001, p = .967, \eta^2 = \text{n/a}$</td>
</tr>
</tbody>
</table>

**7.8.3 Accuracy**

The descriptive statistics in Table 7.18 and Figure 7.3 reveal that the measures for accuracy showed an increase after one semester of study (Japanese 3) for both character-based learners and non-character-based learners. The ratio of EFCs on the pre-test was much higher for non-character-based learners compared to character-based learners, but the rate of development was similar for the two groups.
Table 7.18

*Descriptive Statistics for Character-based and Non-character-based Learners:*

**Accuracy**

<table>
<thead>
<tr>
<th>Measures</th>
<th>Character-based (n = 16)</th>
<th>Non-character-based (n = 25)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test</td>
<td>Post-test</td>
</tr>
<tr>
<td>Ratio of error-free sentences</td>
<td>0.30     0.16</td>
<td>0.34 0.15</td>
</tr>
<tr>
<td>Ratio of EFCs</td>
<td>0.20     0.21</td>
<td>0.34 0.28</td>
</tr>
<tr>
<td>Words per error-free sentences</td>
<td>6.31     2.91</td>
<td>9.44 5.90</td>
</tr>
</tbody>
</table>

**Figure 7.3.** Estimated marginal means for accuracy.

C = character-based learners; NC = non-character-based learners.
Repeated measures ANOVA revealed no statistically significant difference in the change over time between the two groups for all accuracy measures (see Table 7.19).

Table 7.19

Repeated Measures ANOVA Results for Character-based and Non-character-based Learners: Accuracy

<table>
<thead>
<tr>
<th>Measures</th>
<th>Character-based and Non-character-based learners</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratio of error-free sentences</td>
<td>$F(1, 39) = 1.37, p = .248, \eta^2 = n/a$</td>
</tr>
<tr>
<td>Ratio of EFCs</td>
<td>$F(1, 39) = .001, p = .978, \eta^2 = n/a$</td>
</tr>
<tr>
<td>Words per error-free Sentence</td>
<td>$F(1, 39) = 1.06, p = .311, \eta^2 = n/a$</td>
</tr>
</tbody>
</table>

7.8.4 Fluency

As seen in the descriptive statistics in Table 7.20 and Figure 7.4, all the measures for fluency showed improvement from the pre- to post-test, regardless of whether the learners were character-based or non-character-based.

Table 7.20

Descriptive Statistics for Character-based and Non-character-based Learners:

Fluency

<table>
<thead>
<tr>
<th>Measures</th>
<th>Character-based $(n = 16)$</th>
<th>Non-character-based $(n = 25)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test</td>
<td>Post-test</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>No. of words</td>
<td>83.88</td>
<td>34.06</td>
</tr>
<tr>
<td>No. of sentences</td>
<td>9.44</td>
<td>3.86</td>
</tr>
<tr>
<td>No. of clauses</td>
<td>3.94</td>
<td>2.67</td>
</tr>
<tr>
<td>No. of KCs</td>
<td>31.38</td>
<td>18.32</td>
</tr>
<tr>
<td>No. of KWs</td>
<td>18.88</td>
<td>10.06</td>
</tr>
</tbody>
</table>
The results from repeated measures ANOVA revealed that a statistically significant difference existed in the change over time between the two groups only for the number of KCs and KWs, with a large effect size. A post-hoc test pairwise comparison was conducted for these two measures; although both groups improved over time, the character-based learners’ difference in the pre- and post-test mean value was higher than the non-character-based learners’ mean value (see Table 7.22).

Figure 7.4. Estimated marginal means for fluency.
Hence, character-based learners improved more than did the non-character-based learners for the *kanji*-related measures. This suggests that a learner’s background does affect the development of *kanji*-related fluency measures.

Table 7.21

*Repeated Measures ANOVA Results for Character-based and Non-character-based Learners: Fluency*

<table>
<thead>
<tr>
<th>Measures</th>
<th>Character-based and Non-character-based learners</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of words</td>
<td>F(1, 39) = 2.30, p = .138, $\eta^2$ = n/a</td>
</tr>
<tr>
<td>No. of sentences</td>
<td>F(1, 39) = 2.14, p = .151, $\eta^2$ = n/a</td>
</tr>
<tr>
<td>No. of KCs</td>
<td>F(1, 39) = 21.07, p &lt; .001, $\eta^2$ = .35</td>
</tr>
<tr>
<td>No. of KWs</td>
<td>F(1, 39) = 13.80, p &lt; .001, $\eta^2$ = .26</td>
</tr>
<tr>
<td>No. of clauses</td>
<td>F(1, 39) = .117, p = .734, $\eta^2$ = n/a</td>
</tr>
</tbody>
</table>

Table 7.22

*Pairwise Comparison: Mean Difference for No. of KCs and KWs*

<table>
<thead>
<tr>
<th>Measures</th>
<th>Character-based ($n = 16$)</th>
<th>Non-character-based ($n = 25$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of KCs</td>
<td>50.188 (p &lt; .001)</td>
<td>17.840 (p &lt; .001)</td>
</tr>
<tr>
<td>No. of KWs</td>
<td>26.438 (p &lt; .001)</td>
<td>11.480 (p &lt; .001)</td>
</tr>
</tbody>
</table>

7.8.5 Content

Table 7.23 reflects the descriptive statistics for the pre- and post-tests for character-based and non-character-based learners regarding content (i.e., sufficiency and persuasiveness). Table 7.23 and Figure 7.5 show that each measure increased for both groups. Non-character-based learners had a slightly higher initial score compared to the character-based learners.
Table 7.23

Descriptive Statistics for Character-based and Non-character-based Learners:

Content

<table>
<thead>
<tr>
<th>Measures</th>
<th>Character-based</th>
<th>Non-character-based</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 16)</td>
<td>(n = 25)</td>
<td></td>
</tr>
<tr>
<td>Pre-test</td>
<td>Post-test</td>
<td>Pre-test</td>
<td>Post-test</td>
</tr>
<tr>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Sufficiency</td>
<td>1.13</td>
<td>0.72</td>
<td>2.19</td>
</tr>
<tr>
<td>Persuasiveness</td>
<td>1.06</td>
<td>0.44</td>
<td>2.00</td>
</tr>
</tbody>
</table>

![Graphs showing the change in sufficiency and persuasiveness over time for character-based and non-character-based learners.]

a) Sufficiency  
b) Persuasiveness

C = character-based learners; NC = non-character-based learners.

*Figure 7.5. Estimated marginal means for content.*

To determine if a significant difference existed in the change over time between the two groups, repeated measures ANOVAs (between subject factors) were performed. These results reveal that no significant difference existed in the change over time between the two groups for both sufficiency and persuasiveness (see Table 7.24). This suggests that both groups improved over time from the pre- to post-test, but no differences were apparent in the rate of improvement regarding content between the two groups.
Table 7.24

Repeated Measures ANOVA Results for Character-based and Non-character-based Learners: Content

<table>
<thead>
<tr>
<th>Measures</th>
<th>Character-based and non-character-based learners</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sufficiency</td>
<td>F(1, 39) = 0.34, p=.564, η² = n/a</td>
</tr>
<tr>
<td>Persuasiveness</td>
<td>F(1, 39) = 1.00, p = .323, η² = n/a</td>
</tr>
</tbody>
</table>

7.8.6 Structure

The descriptive statistics in Table 7.25 and Figure 7.6 show that both groups improved in coherence and paragraphing from the pre- to post-test. Similar to content (i.e., sufficiency and persuasiveness), non-character-based learners had a higher initial score for both coherence and paragraphing. The rate of improvement was greater for character-based learners (1.06 for coherence and 0.81 for paragraphing) compared to that for non-character-based learners (0.56 for coherence and 0.6 for paragraphing).

Table 7.25

Descriptive Statistics for Character-based and Non-character-based Learners: Structure

<table>
<thead>
<tr>
<th>Measures</th>
<th>Character-based (n = 16)</th>
<th>Non-character-based (n = 25)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test</td>
<td>Post-test</td>
</tr>
<tr>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Coherence</td>
<td>0.94</td>
<td>0.77</td>
</tr>
<tr>
<td>Paragraphing</td>
<td>0.38</td>
<td>0.81</td>
</tr>
</tbody>
</table>
C = character-based learners; NC = non-character-based learners.

*Figure 7.6.* Estimated marginal means for structure

Table 7.26 shows the repeated measures ANOVA results. These indicate that no significant difference was apparent in the change over time between the two groups for both coherence and paragraphing. Similar to content, the results suggest that although both groups improved in coherence and paragraphing over time in Japanese 3, there was no difference in the rate of improvement between the two groups.

Table 7.26

*Repeated Measures ANOVA Results for Character-based and Non-character-based Learners: Structure*

<table>
<thead>
<tr>
<th>Measures</th>
<th>Character-based and Non-character-based learners</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coherence</td>
<td>F(1, 39) = 3.83, p = .058, η² = n/a</td>
</tr>
<tr>
<td>Paragraphing</td>
<td>F(1, 39) = 0.84, p = .365, η² = n/a</td>
</tr>
</tbody>
</table>

**7.9 Summary: Measures Showing Differences in Development**

The above results reveal that differences in development for character-based learners and non-character-based learners were only seen in *kanji*-related areas. The findings also reveal that a difference was already apparent in the pre-test scores between the two groups for *kanji*-related measures. This suggests that a learner’s
background affects development in *kanji*-related measures but not the development of other measures, such as accuracy, content and structure.
Chapter 8: Study 2 Interview Results

Interviews with both students and teachers were conducted in the current study. The two main objectives in conducting student interviews were to: 1) determine whether students’ perception of writing development corresponded to the focus of writing in Japanese 3 and Japanese 4, and 2) examine whether students perceived improvement corroborated the findings from the quantitative analysis (discussed in the previous chapter). The purpose of conducting teacher interviews was to determine whether students’ perceptions of writing development corroborated with teachers’ perceptions.

This chapter will present the findings of the student and teacher interviews. Semi-structured interviews (30-minute duration) were conducted with participants at the end of Japanese 3 and Japanese 4. All interviews were conducted in English. The first section presents the results from the student interviews, while the second section presents the results from the teacher interviews. At the end, a summary of findings from the student and teacher interview results is provided. All quotations presented are direct statements from the participants, using pseudonyms.¹

8.1 Student Interview Results

As mentioned in Chapter 6, in Japanese 3, 40 out of 41 student participants took part in the interviews. In Japanese 4, all 22 students participated in the interviews. Analyses of the transcribed interview responses have been grouped into three broad themes: attitudes to writing in L1 and L2, self-assessment of writing improvement, and factors that influenced the progress in writing. Some themes are

¹ A number of students participated in student interviews for both subjects. Quotations from Japanese 3 and Japanese 4 student interviews are indicated as either ‘Japanese 3’ or ‘Japanese 4’ after each quotation, unless this is stated in the sentence. Due to the very small number of participants for teacher interviews (Hayashi-sensei and Tanaka-sensei in Japanese 3 and Hayashi-sensei in Japanese 4 and follow-up interview), quotations from these participants are not identified further.
only discussed under Japanese 3, while other themes that relate to both Japanese 3 and Japanese 4 will be discussed under both subjects.

8.1.1 Attitudes to writing in L1 and L2

Research on L2 writing has shown that writing proficiency in L1 may affect the way learners write in their L2 (e.g., Cumming, 1989; Zamel, 1983). Recent research has also suggested that interest and self-efficacy determine whether the learner engages in writing (Kormos, 2012). In Japanese 3 interviews, students were therefore asked whether they enjoyed writing in their L1 and in L2 (i.e., Japanese).

More than half of the students (30 out of 40) responded that they enjoyed writing in their L1. However, the majority of those who said they enjoyed writing in their L1 were referring mostly to non-academic writing, such as emails, journals, blogs and stories. Those students who said they enjoyed writing academic essays noted they enjoyed writing more when not being assessed.

Some of the reasons for enjoying writing in their L1 included the enjoyment of creative writing, being able to express thoughts and using a preferred method of communication. For example, not only did Marvella like to write stories, but she also found that by writing to her future self, she could give encouragement to herself:

\[ I \text{ like writing stories. I also like to write blog post so that is like a reflection of my day. I think it is a form of expression and also a good reminder of what happened. Every now and then I write a letter to my future self. Sometimes when you are down and depressed and everything when I read my letters I get encouraged. So it is actually quite good. (Marvella, Japanese 3) } \]
For students like Ben, writing was easier than speaking:

*I find that writing in English, getting my thoughts out is easier than when I am speaking because when I am speaking you have to think what you are saying pretty much as you are saying.* (Ben, Japanese 3)

Out of the 30 students who said that they liked to write in their L1, 28 said that they also liked to write in Japanese. For example, David said that he liked to write in both his L1 and L2. However, with Japanese, he enjoyed writing if it was not assessed:

*Yes, I enjoy writing essays, also enjoy writing in Japanese to an extent but if I had to do it and being marked it gets pretty stressful. If I am doing as a practice like in VCE but that is not marked and you get corrected, I enjoy it. I can write more freely.* (David, Japanese 3)

For most students the reason for liking to write in both languages was similar—they were able to express what they wanted to communicate. For example, Peter’s main reason for liking to write in his L1 was that he could express what he wanted to say in written form more easily compared to in oral form:

*It is an easy form of communication than speaking. It is a much more considered form of communication because it gives you time to edit and think. For me I enjoy writing a lot because it gives me that opportunity to think more to produce something.* (Peter, Japanese 3)

Peter’s reason for liking to write in Japanese was the same as that for writing in English’ however, he noted that writing in Japanese gave him satisfaction as well:

‘I get a sense of achievement when I produce something in Japanese’ (Peter, Japanese 3).
The reason for liking to write in both L1 and L2 was similar for most students, but the reason was different for students like Christina. Christina enjoyed writing in her L1 because it was challenging:

*I like writing in English, I like to write essays because you are trying to prove a point and convince them [i.e., the reader, marker] I like that kind of writing. Get them to see my way, if they agree with me I am pretty happy.* (Christina, Japanese 3)

Her reason for enjoying writing in Japanese was very different from her reason for enjoying writing in L1 (i.e., English)—it was because of how the letters (characters) look, ‘I like to write in Japanese because it is pretty and cute, I like あ because it is all curly and cute’. (Christina, Japanese 3)

For those students who did not enjoy writing in their L1 (10 out of 40), this was because it was time consuming or they could not see the point in writing. Most students who said that they did not enjoy writing in their L1 also said that they did not enjoy writing in their L2. For example, Jimmy’s reason for not enjoying writing was because it took much time for him to write essays:

*Not really because I tend to be very... uum a bit of a perfectionist so when I am like writing essays I will take a long time. How to structure my sentences and the wording like which words to use cause there’s subtle implications of different words.* (Jimmy, Japanese 3)

His reason for not enjoying writing in Japanese differed from his reasons for writing in L1—due to his limited vocabulary in Japanese, he found it difficult to write in that language. He explained that he struggled to write in Japanese and it was even more difficult compared to writing in L1:
I struggle to write in Japanese because I don’t know the language. I just don’t know a lot of vocabularies. I think in English when I write. And then kind of translate in Japanese. So a lot of the times I struggle if I don’t know the vocabulary in Japanese so I can’t translate (Jimmy, Japanese 3)

The reasons given by the 12 respondents (out of 40) who did not enjoy writing in Japanese were similar to Jimmy’s, regardless of whether the student enjoyed writing in their L1 or not. It seemed that many lacked confidence when it came to their L2. They felt they did not know Japanese well enough to write well. It appeared that the level of confidence in their L2 depended on what types of writing they were doing (e.g., simple sentences or writing assignments). For example, Monique (Japanese 3) said, ‘it depends on what I am writing. If I have the tango (words) I guess I am more confident’, and Gabriella (Japanese 3) commented that she experienced difficulties when writing assignments in Japanese, ‘for simple sentences, I can do it, applying all the grammar, that is alright but organising the stuff something for like assignment I cannot’. It was clear from the interview responses that some students were unable to write as much as they wished due to being lower-intermediate learners. For instance, Damien (Japanese 3) commented, ‘I just wish I could write more’, while Alex (Japanese 3) said, ‘with Japanese you have to be able to know a lot to write more. Like if you want to write like a creative story I don’t know how to do that and it becomes limited’.

Surprisingly, some students did not enjoy writing in their own language, but enjoyed writing in Japanese. For example, Helena (Japanese 3) said that she did not enjoy writing in her L1; she had had a bad experience with her VCE English\(^2\): ‘I don’t like writing essays because VCE English was so bad, ohh traumatic. I used to

\(^2\) VCE is a certificate that the majority of the students in Victoria, Australia receive on satisfactory completion of their secondary education.
like writing stories but after VCE I hate it. It ruined it, so not really’. However, she said that she enjoyed writing in Japanese because she wanted to be fluent in another language. Similar to other students, she also mentioned that although she enjoyed writing in Japanese, she struggled: ‘Mainly remembering kanji and expressing things in Japanese is difficult for me’.

The student interviews suggest that those who enjoyed writing in their L1 also tended to enjoy writing in Japanese. Writing in Japanese, however, gave them a different kind of enjoyment compared to writing in their L1 e.g., in a sense of self-achievement. In contrast, those students who said that they did not enjoy writing in their L1 struggled even more in Japanese, as they felt they did not know Japanese well enough (e.g., a lack of vocabulary) to express their ideas.

8.1.2 Self-assessment of writing improvement

Having an understanding of participants’ perspectives on their writing development through self-assessment is important, as it may reveal not only whether some quantitative measures are not capturing aspects of development, as well as some of the factors that lead to perceived improvement or that may explain the lack of improvement.

In both Japanese 3 and Japanese 4, students were asked whether they felt that their writing in Japanese had improved over the semester. In Japanese 3, the majority of students (35 out of 40: 88%) responded that they had improved. Of those students who continued on to Japanese 4 (22), most (16 out of 22: 73 %) thought that their writing improved during that course compared to Japanese 3. The main areas the students felt they had improved in were grammar (accuracy and sentence structure), vocabulary, and kanji (each of these is discussed below). Other than the accuracy,
students’ perceptions of writing improvement corroborated the quantitative findings presented in the previous chapter (i.e., improvements in structure, fluency and kanji).

8.1.2.1 Grammar improvement

Grammar was one of the most frequently reported areas of improvement. Half of the students (20 out of 40) thought that their grammar improved in Japanese 3, and a majority of students (14 out of 22) also felt that their grammar improved during Japanese 4. Eleven students thought that their grammar improved in both subjects. Most students in both subjects felt they could apply what they referred to as the ‘grammar patterns’ (sentence patterns with a set of structural rules governing the composition of clauses, phrases, and words) they had learned during the semester. For example, in Japanese 3 Damien mentioned that he only had to learn a few grammar patterns, which he could remember and was able to apply in his writing:

"I think this semester in Uni it is more about using it in real life which I think it was good, so I think I improved more this semester than in the year of VCE…. The grammar patterns are actually sticking in my mind…. We only learn like two or three grammar patterns a week, this is kind of study that I like. Because there is only a few I can concentrate. (Damien, Japanese 3)"

In Japanese 4, Damien also felt that his grammar improved mainly through more practice, ‘Yeah with more grammar patterns and more practice with writing’. Matt also mentioned in Japanese 3 that having exposure to different grammar patterns during the semester ensured he improved his writing, ‘understanding the grammar patterns. Learning the new grammar patterns were the major differences I think’. In Japanese 4, Matt explained that he took a different approach to writing compared to the previous semester in which he thought improved his sentence
structure, ‘yes [p]robably incorporating a more holistic approach to the task. More
grammar patterns and planning the whole essay’.

However, for some students, applying the grammar they had learned was a
little difficult. For example, Nikki thought that her grammar improved in Japanese 3
but she was not very confident: ‘grammar maybe… how to make sentences. I know
grammar but sometimes I cannot use it’. However, in Japanese 4 she was more
confident in applying the grammar patterns that she learned: ‘I learnt more grammar
points which are the most important thing because we need to connect sentences’
and she felt this assisted her writing.

8.1.2.2 Vocabulary

Students also felt they improved in the area of vocabulary. Several students
(12 out of 40) in Japanese 3 commented that they had more vocabulary by the end of
the semester. In Japanese 4, a few students (5 out of 22) also felt they had more
vocabulary after the semester.

By having more vocabulary, Luke (Japanese 3), for example, noted that he
was able to think more in Japanese when studying. He commented that compared to
his pre-test, he thought more in Japanese when doing the post-test, ‘I had more
vocabulary. I was thinking more in Japanese. A lot of spelling mistakes though.’ He
also felt that he improved in vocabulary at the end of Japanese 4, ‘vocabulary, so
many more words that I know’. Luke only stated he had more vocabulary after one
semester, but students like Mary mentioned that by increasing the amount of
vocabulary her expression improved as well. In both subjects, Mary commented that
she learned a lot of vocabulary, which was one reason she felt her writing improved.
In Japanese 4, she said that by further building up the vocabulary from Japanese 3, it
assisted in her writing, ‘I think I improved on the writing, yes. I can express more. Last time I couldn’t express because of the limited vocab’.

Having a wider vocabulary also helped students to write longer texts. In both Japanese 3 and Japanese 4 interviews, students such as Yasmine and Norma gave similar reasons for improvements in their writing. Norma (Japanese 4) said she felt that she was able to write longer passages in her post-test compared to the first time she wrote (i.e., the pre-test) as she had more vocabulary, ‘I think I have more vocabulary and I can write longer’. Yasmine (Japanese 3) also commented that she was able to write longer passages for similar reasons: ‘some of the words, I didn’t know how to write it, before I can only speak it’. She explained that she watched Japanese drama frequently and had experience of Japanese outside class, but she was unsure of how to write some words that she had acquired through watching Japanese drama.

8.1.2.3 Kanji improvement

Another frequently reported area of improvement was the command of kanji. More than half of the students (28 out of 40) in Japanese 3 felt that they improved in their kanji use. A similar proportion (12 out of 22) felt that their kanji improved in Japanese 4 and that they could incorporate them in their writing. For example, Samantha (Japanese 4) said, ‘yeah, like I can use more kanji in my writing’. When students used more kanji instead of hiragana or katakana (the other two Japanese scripts) in their writing, they felt they had improved. As Candice (Japanese 4) commented, ‘yes, yes, so I don’t have to use so many katakanas, I can use more kanji’. Moreover, knowing more kanji allowed students like Alice (Japanese 4) to save time as she did not have to look up the kanji (in the dictionary): ‘I think it was
easier for me to write because I didn’t have to go and look up the kanji, especially when handwriting’.

Although many students felt they improved in kanji, learning kanji was clearly difficult for some students. In both subjects, students had weekly kanji tests (non-assessed) and those from character-based languages had some advantage. In Japanese 3, students were asked how long it took them to learn the required ten KCs every week. The average time (minutes) taken by character-based learners to learn 10 kanji (15 minutes) was less compared to that for non-character-based learners (65 minutes).

Even though the results show that character-based learners had an advantage over non-character-based learners, they did not necessarily feel that kanji was easy: in many circumstances, they were unsure whether they were using the right kanji or mistakenly using Chinese characters. For example, Samantha (Japanese 4) is a character-based learner and in Japanese 3 she commented that she often confused kanji with Chinese characters. However, having the weekly kanji test helped her distinguish some of the differences and in Japanese 4 she was more confident with her use of kanji: ‘I think it is good that we have kanji dictation in Seminar 1 so that I can know which kanji I machigai [make mistake]’. Jasmine, also a character-based learner, felt the same way as Samantha after Japanese 4, ‘kanji got easier….after this semester I can really tell the difference, before I thought it was kind of the same but couldn’t really tell the difference. Melissa (Japanese 3), who is a non-character-based learner, said that she enjoyed learning kanji and did not struggle to memorise them. The fact that she was also learning Chinese as a FL made kanji interesting for her, ‘yeah, I enjoy it. I think it is one of my strength. And also because I have to memorise Chinese as well’. In Japanese 4, she also felt that her kanji got better:
Kanji I would have improved. Learning Chinese has helped me improve my kanji. I remember when I was learning kanji in high school, it would be really ugly and I wouldn’t have the order right but I am a lot better at it now. (Melissa, Japanese 4)

Melissa’s comments also suggest that students were not only putting effort into memorising kanji, but they were also concerned with how the kanji looked when they wrote them. For example, Christina, who is also a non-character-based learner, felt that she improved in her kanji because she was able to write them better.

Enjoyment in writing kanji and a sense of improvement were not always strongly related. Students felt they had improved in their kanji regardless of whether they enjoyed learning kanji. About one-third of the students in Japanese 3 (15) who felt that they had improved in their kanji also commented that they enjoyed learning kanji. Some enjoyed kanji as they looked like pictures rather than words. Christina (Japanese 3) said, ‘I find it interesting, trying to make picture out of it. I tried to make a picture with all of them, yes I enjoy it’. It is interesting to note that those who said they enjoyed learning kanji were all non-character-based learners (character-based learners neither enjoyed nor disliked learning kanji). For some of the non-character-based learners, although it took time to learn kanji, it was something they enjoyed. However, for others it became a struggle. For example, Corinna (Japanese 3) spent five hours a week to learn the required 10 kanji, ‘five focused hours throughout the week, yes. Or it may take more. I guess it depends. Each kanji 25 times’. Even though it took her a long time and she found them difficult to remember, she still found it interesting:
I enjoy it. I find memorising the kanji itself is the tricky part because you have to associate it with the reading and remember it day to day you have to be using it. (Corinna, Japanese 3).

Some students improved in their kanji even though they did not enjoy learning kanji. For example, Alex (Japanese 3) commented that although he did not enjoy learning kanji during the semester, he learned them throughout the semester because students had to report on their kanji test mark (not assessed) in front of everyone, ‘I don’t enjoy it but it is not difficult. Sensei actually asked for the marks so probably that is why people put in effort’. George, a non-character-based learner, also did not enjoy learning kanji because of the time it took to learn them:

Because of my background, it takes a bit of time but I put my time so it is not a struggle for me. I don’t think I enjoy it but it is not terrible. But it is terrifying how many kanji you need to know to be able to read a newspaper. (George, Japanese 3)

In Japanese 4, although George felt that building up the number of KCs over two subjects through the weekly test helped him improve his writing, on some occasions he found it difficult to learn the required KCs for the week:

I think so. I think it did help but can be daunting sometimes because when you have other things to do and you realise two days before that you have to learn eight kanji, you just hope that they are the easy ones. (George, Japanese 4)

Most participants felt that their writing had improved, especially in grammar, vocabulary and kanji. In both subjects, more students noted improvements in grammar and kanji compared to improvements in vocabulary. The factors that led to
these improvements, as well as the factors that hindered further progress in students’ writing, are discussed in the following section.

8.1.3 Factors that influenced writing development

Students felt that several factors had led to their writing development, or the lack thereof. These factors were practice, reading and feedback (teacher feedback and peer feedback).

8.1.3.1 Practice in writing

One factor that the students perceived as important for improving their writing was practice. The amount of practice that learners undertake may depend on the assessment tasks and classroom activities. In both Japanese 3 and Japanese 4, assessment tasks that included writing were individual assignments, group assignments and writing a script for oral assessment. Classroom activities that included writing were weekly kanji tests, a weekly reflective writing task, and textbook questions. Assessment tasks are discussed first, followed by classroom activities.

Individual assignments in Japanese 3 required students to produce a tourist webpage in Japanese (600 words). In Japanese 4, the research task required students to write on any topic covered in lectures during the semester (600 words). In both subjects, approximately half of the students commented that individual assignments were either useful in improving their writing or that when producing them, they attempted to use grammar/vocabulary/sentences they had learned during the semester. Therefore, students felt that the assessment tasks gave them the opportunity to practice their writing. In Japanese 3, Jimmy commented that it was necessary to revise all the grammar he had learned during the semester to write his individual assignment:
The final assignment was good because that was covering the whole semester so you had to go back and learn all the grammar points and find a way to include that to the assignment. (Jimmy, Japanese 3)

Alicia, who studied in both Japanese 3 and Japanese 4, found the individual assignments allowed her to practice various grammar and verbs:

The individual assignment was very good for that [writing improvement]...

You repeat all the grammar and it is good. (Alicia, Japanese 3)

I guess it was pretty hard to write but I really liked it because it kind of made me use a lot of verbs and that was quite enjoyable. (Alicia, Japanese 4)

Most students found the individual assignment difficult compared to the group assignments, but at the same time, they felt they gained more writing practice, especially in terms of using new grammar (sentence patterns) and vocabularies. George, who studied in both semesters, found the individual assignments challenging. In Japanese 3, he explained that re-writing the information that he obtained in his own words was difficult. Moreover, he tried to incorporate the techniques he had learned during the semester in his assignment:

It was definitely good for my Japanese. It pushed me. I found a bit stressful writing. Making sure I didn’t copy it. I wanted to understand and write it in my own words. I also wanted to use the Japanese that I learnt. Converting the stuff that I read into my own words was very difficult. (George, Japanese 3)

In Japanese 4, he explained that some of the grammar patterns from the resources he obtained were difficult to use. He did attempt to use the unfamiliar grammar patterns: ‘I often didn’t want to look at the grammar that they used because I didn’t want to plagiarise at all and when I looked at mine, it looked nothing like it’.

233
Using and practicing the grammar he had learned over the semester in class was better than using the grammar patterns he obtained from the resources: ‘I thought about it and I was happy that I did it that way….The grammar that we learnt this semester was the main help’.

In comparison with the individual assignments, it appeared that the group assignments helped students more with their reading rather than with their writing skills. Students reported they had no opportunity to practice Japanese writing in the group assignments, as most of the group assignments in both subjects were written in English. In Japanese 3, the group project consisted of three stages: 1) write an itinerary (in Japanese) from Narita airport to Amanohashidate, a tourist destination in Japan; 2) write a discussion piece on the destination (in English); and 3) write an annotated bibliography (in English). Out of the three stages, only Stage 1 required students to write in Japanese and as Monique (Japanese 3) pointed out, even that writing was brief: ‘the only thing that we needed to do in Japanese was the first one (stage 1), and most of it was in point form’.

Group assignments for Japanese 4 were designed to prepare students for their individual assignment. Two group tasks were required, both in English. The first one required students to discuss and write about their topic and the second one required them to write annotations for their bibliography. The group work may have helped students gather information but, as the majority of students (18 out of 22) pointed out, it did not help them improve their writing in Japanese as they were written in English. For example, George (Japanese 4) believed that the group assignment was only useful for gathering information but not for improving writing, ‘I don’t think it helped me with my writing as such. It was good to work with other people in terms of what they found and show each other’.
The resources that students used for the group assignments were in Japanese; therefore, many students found that their reading skills developed, but not their writing skills. As Monique said:

*It was mostly in English so it didn’t help that much. It was mainly talking with the group members about it. The phase 2 was the bibliography. So I think it helped a lot with the reading. Like after this semester, I can read Japanese a lot better but I don’t know about writing because you have to think how to form a sentence yourself.* (Monique, Japanese 4)

Other than the individual and group assignments, an oral assessment was required in both subjects. Students had to create a 5-minute video in a group of three. In order to do so, students had to write scripts. Although the script was not assessed in either subject (only the video was assessed as an oral assessment), more than half of the students (25/40 in Japanese 3 and 20/22 in Japanese 4) said that it helped with their writing. Sara commented in Japanese 3 that she was able to revise and practice her grammar while writing the script with her peers, ‘yes—one of our group members made a list of all the grammar that we learnt and we just checked whether we used them all’. In Japanese 4, Sara mentioned that this task allowed her to learn from her group members again and gave her writing practice as she was required to write part of the script, ‘we all did work on each part. Divide our script into three different things and each one of us wrote on one thing’ (Japanese 4). Students like Mary (Japanese 3) could practice writing during the script writing process as they had to edit a number of times, ‘it helped me a lot—with the video you need to be perfect when you speak so you keep editing and editing for it to make it better so it helped me a lot with writing’. However, in Japanese 4 Mary did not feel the same way, mainly because she did not have the opportunity to write the script:
It was only the leader (each group consisted of three students with one person being the leader) that did it. She just like said, memorise this script and we were like... ‘ok’. It didn’t help me with my writing at all. Last semester was really good because we did it together sentence by sentence but this semester. (Mary, Japanese 4) 

Some groups had students with a slightly higher Japanese proficiency than others. In those groups, the higher-proficiency level students did more writing. For example, Melissa commented that although she did have the opportunity to write a part of the script, she wrote less than did the others:

*There were people stronger in the group than me and you can rely on them to do it. When we divided the work up there was less for me to do and I didn’t have to do much. It wasn’t as challenging, the stuff that I had to do. It was like I could use very easy vocabulary, the vocabulary that I already knew and easy grammar stuff. (Melissa, Japanese 3)*

Students like David did not feel that this task was beneficial in Japanese 3, but felt it was useful in Japanese 4. In Japanese 3, he wrote the script in English and other members of their group translated the script into Japanese. For example, David explained how his group worked in making the script for the video:

*I wrote the English script and wrote suggestions on grammar what sort of stuff we could use and someone else wrote the Japanese and the other person edited it. Yeah I mainly wrote the English so I didn’t practice writing in Japanese. (David, Japanese 3)*

In Japanese 4 however, David mentioned that he benefited from this task, as he had to write one section of the script, ‘yeah we divided our work… I wrote one
out and we corrected each other’s work. Yeah, I guess any kind of writing practice help you improve’.

Classroom activities that included writing were weekly activities from the textbook, a weekly kanji test and weekly reflective writing tasks (3 sentences). Each activity will be discussed below in terms of whether it provided practice in writing, and opportunities to learn specific aspects such as kanji and grammar.

The majority of students (in both subjects) answered that the weekly activities from the textbook (mainly grammar practice but also occasional readings) assisted in their writing. Damien suggested that questions from the textbook were useful as they allowed him to practice what he had learned:

*The questions from Genki [Genki is the name of the textbook used in Japanese 3 and Japanese 4] were good because you had to read the book and you had to understand what it was saying and apply to your writing. Try to make the sentences as long as possible. Yeah, that helped me with my writing more than the three sentences each week.* (Damien, Japanese 3)

In Japanese 4, Damien also mentioned that grammar practice in class was useful for his writing, ‘in class, seminar one was grammar so it improved writing by learning more grammar points’. Jasmine also commented in Japanese 4 how the weekly grammar practice had assisted her writing:

*I think we did a series of sentences… comparing sentences that was for grammar that was super similar. That was good in terms of thinking about grammar and how to express.* (Jasmine, Japanese 3)

By later revising the grammar she had found confusing in class clarified it for her and enhanced her practice to use them appropriately.
During Japanese 3 and Japanese 4, students learned eight to ten KCs every week. They were tested on these (non-assessed weekly kanji test). In Japanese 3, a few students commented that the weekly kanji tests helped them with their writing. The majority of students in Japanese 4 (17 out of 22) also felt that the kanji test helped them with their writing. The kanji test not only forced students to learn 10 new KCs, but because it was a dictation test (including the new KCs), it allowed students to learn new sentences and practice writing. In both subjects, Norma thought that the kanji test was useful. In Japanese 3 she said, ‘yes, the kanji test was useful because the teacher was reading the sentences, I have to first guess the meaning of the sentences’. She further elaborated in Japanese 4, ‘I can incorporate more kanji in my writing… I can remember because we have quiz and it forces me to remember the kanji’. This indicates that by having the kanji tests weekly in both subjects, students can use the kanji they learn in their writing. Moreover, in Japanese 4, students were also being dictated kanji that they had learned in the previous semester, which helped them revise and practice. As Alicia noted:

\textit{Actually I think I did better this semester because in the test he was asking this week’s and last week’s. Last year’s even. The kanjis. So it was really interesting because I needed to remember a lot of old kanjis and it kind of helped me to recall. (Alicia, Japanese 4)}

However, even though students learned their kanji every week, many said that they were difficult to remember. For example, Daniel said:

\textit{I study hard for week 1 and once it gets to week 3 I completely forget it. And it becomes really hard. And eventually when I have exams, I have to remember from week 1 – week 12 again. (Daniel, Japanese 4)}
Many character-based learners found the kanji test helped them remember some of the differences between the Chinese characters and kanji, but it seemed that the test did not have a huge impact on improving their writing. Mary (Japanese 4) commented, ‘I already know the kanji. It didn’t really help me’.

In both subjects, insufficient practice was the major factor that most students identified as hindering further progress in their writing. The main form of in-class writing practice observed in both semesters was the weekly reflective writing (non-assessed). They were required to write three sentences reflecting on what they had learned during that week. Students completed this reflective writing task at the end of the class in both subjects and had 10 minutes (sometimes only five minutes) to complete it. Even though this task was conducted every week, it was only three sentences and students did not feel they were getting enough practice. Moreover, although it was meant to be a weekly writing practice, due to the limited class time, this task was not conducted regularly during Japanese 4, ‘They didn’t check like Semester 1 so I only did it on days when they reminded us’ (Monique, Japanese 4).

Matt commented how much more had he practiced when in secondary school:

*To be honest, they were short and brief. In high school we had to write about 400ji [ji is equivalent to ‘words’ in English]. In VCE we had to do writing exam, I found that useful. 3 sentences were a bit brief to get feedback. Even if 200ji it might be alright. (Matt, Japanese 3)*

As students did not have much opportunity to practice writing in class, students like Anna felt that it was very difficult for her to write an individual assignment:
Yeah.. I just feel like we don’t write a lot. We do like grammar but that is just writing out sentences. Sensei expects us to write an 800 word essay but we don’t do much in class, so it is really difficult transition. (Anna, Japanese 4)

In addition to being brief, the majority of students in Japanese 3 (34 out of 40) said that the instructions for the weekly reflective writing task in class were very vague. It was unclear what kind of writing practice was required. For example, Helena (Japanese 3) said, ‘if they had clear instructions it would have been better. Most of the time, I didn’t know what to write’.

More than half of the students said that they enjoyed writing in Japanese (see Section 8.1.1), but it seemed that most did not practice their writing outside class. The majority of students watched anime or drama outside class, but practicing writing was difficult. Some suggested that they should have been given a task to practice on their own. Edward (Japanese 3) said, ‘if there is something to write about, if they give us something I will write’. Students were aware that it was necessary to practice. As Melissa (Japanese 3) said, ‘unless I have a test or something I won’t write it. I think that is the nature of University. I should practice more’.

A few students commented that they practiced outside class using the textbook, the workbook (a separate workbook they can buy based on the textbook) or writing a diary:

I like to use the workbook. I always found it hard when they had broad question in the workbook. Going through the examples and writing was good.

(George, Japanese 3)
I have started to write a Japanese diary. I think it helps, sometimes when I am walking the street and if I see something interesting I will put it in my diary. I am constantly trying to think in Japanese. (Anne, Japanese 3)

George, who continued on to Japanese 4, kept using the workbook to practice. Anne was keen on studying Japanese, but as she was unable to take more elective subjects, she could not continue after Japanese 3.

Others, like Joanna (Japanese 3), messaged her Japanese friends frequently, which she felt helped her writing, ‘I have a few friends that I message regularly. I find it is really good practice’.

8.1.3.2 Reading in Japanese

Past research (e.g., Lee & Hsu, 2009; Mermelstein, 2015; Tsang, 1996) suggests that reading has a positive impact on writing development. In the current study, more than half the students in Japanese 3 and Japanese 4 also reported that reading benefited their writing. In both subjects, students were required to use Japanese resources (e.g., websites, articles, books) to do their group and individual assignments. Students said that reading Japanese resources helped with their writing mainly because the Japanese resources provided examples (i.e., grammar patterns and sentence structure) that they could use in their writing. For example, Anna (Japanese 3) mentioned that, ‘by reading the websites I think it helped with ordering the sentences and also with the particles’. And in Japanese 4:

Sometimes you will see a grammar pattern in a sentence that I would have thought I would be using and things like that. And also how setting everything out, like particles and things like that like make the entire sentence. (Anna, Japanese 4)
Other students mentioned that reading helped with learning vocabulary and information about their topic. In Japanese 3, Yasmine noted that students were able to learn a new vocabulary related to their assignments through reading Japanese websites:

Yes, especially because we did Izumo-taisha [Izumo-taisha is a name of the shrine that Yasmine’s group chose as one of the destinations] and we got a lot of special words. (Yasmine, Japanese 3)

In Japanese 4, Yasmine’s comments were more about having an understanding of the topic (her topic was on kanji):

It helps a lot because it tells you a lot about kanji and something like that. I figured Japanese system comes from Chinese system but I figured totally wrong. So the resources helps us understand in a new way. (Yasmine, Japanese 4)

Some students, such as Matt, felt that having to re-phrase the materials into their own words helped them improve their expression in Japanese:

Reading Japanese sources and re-wording into simplified Japanese words to put it I guess into context, this semester was a website, but you can also do that in other contexts such as journals. By reading Japanese and re-writing into your own words – I felt I improved in expression. (Matt, Japanese 3)

In Japanese 4, Matt thought that the resources gave him more understanding about his topic. In addition, they provided examples of how academic essays should be written, ‘I guess we learnt more about the topic that we were doing. I found them beneficial for future academic essays’. (Matt, Japanese 4).

Reading Japanese websites also allowed students to confirm whether their expression was correct or not, as illustrated by Joanna’s comment:
It does give good examples of what is correct. If I wrote one thing and if the Japanese websites have written it different way it does give me a second guess that my work is not correct, so it is a good reference to know what is commonly used in Japanese. (Joanna, Japanese 3)

For those students who thought that the resources did not help with their writing, this was because the resources were too difficult for them to read. Most of the Japanese resources were for native Japanese speakers (i.e., not for L2 Japanese learners) and therefore consisted of too many kanji that students could not read. In order to understand the content, most students said they used translations from Google.

8.1.3.3 Feedback

Another factor mentioned by the students’ as affecting their writing development was feedback. The literature on feedback for writing mentions two types of feedback: teacher and peer. Empirical studies on teacher feedback (e.g., Bitchener, 2008; Bitchener & Knoch, 2008, 2009, 2010; Fazio, 2001) indicate that teacher feedback (especially error correction) is beneficial in improving students’ writing accuracy. Peer feedback can also assist students’ writing if adequate peer review training is provided (Min 2005, 2006; Rahimi, 2013; Zhu 1995, 2001). In both Japanese 3 and Japanese 4, students received feedback from their teachers and peers. Teacher feedback will be discussed first, followed by peer feedback.

8.1.3.3.1 Teacher feedback

Students’ perception of teacher feedback is important (e.g., Diab, 2005, 2006) as it may explain how and why learners respond to feedback. It is thus important to understand whether students feel they are receiving enough feedback, and what kind of feedback they prefer and find useful. In the current study, students
received feedback on their writing from their teachers for the weekly writing tasks (discussed below) and individual assignments (they only received marks for group assignments). The individual assignments in both subjects were returned after the semester finished and therefore, the researcher could not ask students about these in the interviews.

One reason students gave in relation to difficulties in improving their writing in Japanese 3 was that they did not receive enough feedback on the weekly writing tasks. The reflective writing task was conducted at the end of the class. From the classroom observation, it was apparent that in most cases teachers did not have enough time to give feedback to every student. Most students (24 out of 40) said that they received little or no feedback and attributed this to the lack of time:

Sensei comes around and checks briefly, sometimes sensei doesn’t have time.

I ask for feedback and that was helpful. (Daniel, Japanese 3)

Sensei kind of ticked, so if I get it every week it would have been helpful.

(Luke, Japanese 3)

The above comments from students suggest that when and if they did receive feedback from teachers, they found it useful. The weekly writing task was not assessed, so in some weeks teachers did not give feedback at all, as Peter mentioned:

Initially sensei went around and read through and corrected it. There were correction marks—but towards the end because having going through a lot of materials in the seminar it got less important. At the same time it was important because if people were writing and there was no feedback at all...

that little constant feedback is good. (Peter, Japanese 3)

Those students who said that they received feedback seemed satisfied with feedback from the teachers:
Yep, sensei came over and corrected and also gave us suggestions on better way to say it. Yep they were helpful. (Edward, Japanese 3)

Yeah I appreciate when they correct mistakes, a lot of the times when you are speaking, in casual conversation, people won’t correct you, it is nice when they look at the text and tell me which particles are wrong, things like that because I struggle with those. (Joanna, Japanese 3)

Mary, who studied both subjects, was not satisfied with the teacher feedback in Japanese 3, but was satisfied in Japanese 4. In Japanese 3, she said that the teacher’s comments often just indicated that the grammar was wrong but did not specify the correct form. This type of feedback de-motivated her:

Yeah, sensei [means ‘teacher’ in Japanese] would just nod his head or say that this one is wrong but he didn’t say you should use this to improve or things like that. If he gave me feedback, if he circled this and wrote what is wrong, or if he said you can replace with this word… that would be really useful but because he didn’t do that, it is a bit de-motivating. (Mary, Japanese 3)

In Japanese 4, Mary made an effort to go and ask the teacher for feedback when she was unsure. She commented:

Each week we learnt different vocabs and grammar so when I use it to create a sentence I am not too sure whether it is right or wrong… Technically I go up to her and she will help me… And sensei can tell me whether that is the correct way of writing it. (Mary, Japanese 4)

Overall, participants in the current study felt that teacher feedback was important for writing development and it would have been beneficial if they had
received more feedback from their teachers. The following section outlines what students perceive about peer feedback.

8.1.3.3.2 Peer feedback

During Japanese 3 and Japanese 4, students had the opportunity to provide peer feedback while working on their group tasks. The peer feedback for these tasks was not assessed. Japanese 4 included a minor assessed peer feedback task in which students were required to provide feedback on the draft they wrote for their individual assignment. Non-assessed peer feedback (related to group work) will be discussed first, followed by assessed peer feedback.

8.1.3.3.3 Non-assessed peer feedback (group work)

During Japanese 3 and Japanese 4, students had the opportunity to receive feedback from their peers while working on their assessment tasks. Students indicated that group assignments in Japanese 3 and writing a script for an oral assessment (in both subjects) allowed them to discuss and receive feedback. In Japanese 4, students also had the opportunity to receive peer feedback on their weekly writing tasks in class.

As mentioned previously (see Section 8.1.3.1), one of the group assignments in Japanese 3 was to be written in Japanese (all the group assignments in Japanese 4 were to be written in English). Over half the students (23 out of 40) in Japanese 3 thought that this group assignment was beneficial for their writing. The majority (20 out of 23) thought it helped with their writing because of the peer feedback. Sometimes students were not aware of their errors and through being corrected by their peers, they became more aware of their grammatical errors. Peter’s comment illustrates how being corrected by his group members helped him with his writing, as it made him aware of errors he could not have picked up had he worked on his own:
We were using Google Docs, and we would be working on different sections of the document and write at the same time and we would go through it and review it together. A lot of grammar mistakes... they picked it up. Definitely made the writing that I produced better when reviewed as a group. (Peter, Japanese 3)

He further reiterated the importance of peer feedback when he was commenting about his individual assignment:

When you are in a group there are other people to validate and say that this is good but when you write by yourself, you only have yourself and you are less sure what you have written. (Peter, Japanese 3)

The advantage of peer feedback can also be seen in Corinna’s (Japanese 3) comment, ‘we found that having the extra pair of eyes was actually something helpful…rather than scanning it through ourselves’.

Some groups contained students with a slightly higher Japanese proficiency than others and they could correct other students’ writing, as mentioned by Yasmine and Melissa:

There was a girl in our group that was quite good at Japanese, she was better than the other two. Sometimes when we make mistakes she will correct our work. (Yasmine, Japanese 3).

There were people stronger than me... If I made mistakes in my writing they would fix it so it was good because I knew what I did wrong (Melissa, Japanese 3).

Students like Yasmine and Melissa thought that being corrected by their peers improved their writing.
Another opportunity for non-assessed peer feedback was in the oral assessment tasks. As mentioned previously, in both subjects students had to write a script for their oral assessment task (create a 5-minute video). Although the script was not assessed in either subject, most students thought that this task contributed to their writing development as it allowed them to correct each other’s work and learn from each other. In both subjects, Anna felt that writing the script together as a group and obtaining other people’s opinions allowed her to construct sentences better. In Japanese 3, she said:

*We wrote most of it together, we met up. It was good because some people had different ideas. Yeah and just to help structure the sentence and get other people’s opinions (Anna, Japanese 3)*

In Japanese 4, Anna’s group used Google Docs and corrected each other’s work while writing the script together (through the computer):

*We did ours in Google Doc so everyone was writing and correcting, yeah so it was good because we could correct each other’s work. If we didn’t know something, we could get someone to help us out… just having that feedback helps. (Anna, Japanese 4)*

Monique also commented in both subjects that writing the script was helpful, especially for grammar as she received suggestions about using grammar from her peers:

*When I use a certain grammar some will say that you can use other grammar. They will pick it up. Helped me with grammar. (Monique, Japanese 3)*

*They were able to give feedback, like you can also use this grammar. (Monique, Japanese 4)*
It is interesting to note that students perceived peer feedback as useful when they were working on their group assignments, but did not find it as beneficial when they received feedback on their weekly writing tasks during class. As mentioned previously, in Japanese 4 the weekly writing tasks done in class consisted of only three sentences and were not assessed. Students only received peer feedback on these tasks.

The majority (14 out of 22) of students thought that the feedback received from peers on this task was not helpful. For example George (Japanese 4) said, ‘they were just simple things, like a kanji was wrong and it is something I am used to (kanji being wrong). So I didn’t find it useful that much’. Alicia (Japanese 4) also commented that she did not receive a lot of feedback, ‘they just said you have this kanji or that kanji wrong and that’s it, they didn’t give much. It wasn’t that helpful’. Moreover, it seems that the peer feedback activity was not done regularly. As Anna (Japanese 4) noted, ‘uum, it wasn’t very consistent, I think I only did it a couple of times’.

Some students also admitted that they were either not confident enough or felt reluctant to give feedback. Therefore, the amount of feedback they provided on these weekly writing tasks was limited. Moreover, when receiving comments, students did not trust their peers’ comments, as their peers were at the same proficiency level. For example, Melissa said that she was hesitant to give feedback to peers as she was not confident enough:

*I think it was hard for me to give feedback because my Japanese is not that good. So I didn’t want to be like “you got this grammar wrong”, I’d rather say, I don’t know whether that is right, you better check with the teacher.*

*(Melissa, Japanese 4)*
In terms of receiving feedback from peers, Melissa was also unsure how much of the peer feedback she should be taking into account:

Yeah I got feedback but when someone says ‘oh you shouldn’t do that’ but I wouldn’t be just be like ‘ok’ you know, I will check with the teacher cause they may be wrong you know, they have only been studying Japanese for 2 years. (Melissa, Japanese 4)

Similar to Melissa, Daniel (Japanese 4) also commented that it was difficult to provide feedback to his peers because he could not remember all the grammar he had learned, ‘the grammar that I learnt during that week, I could help but if it is next week I can’t because I will forget’.

Further, as the feedback was not assessed, some students were not enthusiastic about giving feedback. For example, Mary (Japanese 4) explained that she usually sat with her group members (i.e., the group for the assignment) who were not enthusiastic, ‘no I didn’t do this because my group wasn’t good and they didn’t want to do it’.

The findings suggest that students’ perception of peer feedback differed when they were working on their group work and weekly writing tasks. Nevertheless, it is important to note that although the peer feedback itself was not assessed in both cases, the tasks for the group work were assessed but the weekly writing tasks were not assessed.

The findings suggest that students had mixed attitudes towards peer feedback; their attitudes depended on the task. One consideration for these differences may be related to the assessment, whether the task itself was assessed or not.
8.1.3.3.4 Assessed peer feedback

In contrast to the group assignment, oral assessment and weekly writing tasks, a peer review of a draft of the individual assignments was an assessment task in Japanese 4 (5% of the total mark), which the peer feedback was also assessed. Peer review was done online and anonymously. Each student was required to give feedback on two assignments and receive feedback from two students on their assignments.

Almost all (19) students indicated that giving feedback on the individual assignment helped with their writing. The main reason for this was that when students read their peers’ work they were able to reflect on their own writing. When they discovered other students’ errors, they tried not to make the same errors. For example, Cindy (Japanese 4) said, ‘when I was commenting on other people’s structure, I will think about my own essay, what did I do and what I didn’t do’. Nikki (Japanese 4) was not only able to find her mistakes by reading other people’s writing, but also found their writing better than her writing, which gave her examples of how to write well, ‘I found out that there was a lot of things that I missed by looking at others. Also sometimes their sentences are very good, very clear’.

Most students (14) said that the feedback they received from their peers helped them write better for their final individual assignment. However, they also indicated that this depended upon who provided the feedback, as pointed out by Melissa:

*One in particular helped me a lot because it told me this is wrong grammatically, this is wrong, you should change your layout in this way. Other one was not as helpful as that one. Obviously it depends how well the*
person marks on your assignment yeah, I thought it was helpful. (Melissa, Japanese 4)

Sometimes positive peer feedback was misleading. Students who received only good comments from their peers and thought that their draft was correct were then surprised to find that upon submission they received negative comments from their teacher.

For example, Anna and Yasmine mentioned:

Both of them said they were good and I didn’t correct much but then sensei was like you don’t have enough details and all this stuff and I was like...? (Anna, Japanese 4)

One of them, she just said I had one thing wrong with the kanji so I thought it was good but when I got my marks back it was like ‘oh my god’ I did so bad. But I guess that is the disadvantage of peer review because they are not natives. (Yasmine, Japanese 4)

Students’ perceptions of feedback seemed unchanged whether the peer feedback was non-assessed or assessed. When they were given useful feedback, students thought that the feedback was beneficial; however, given that students were still at the lower-intermediate level, many were reluctant to trust their peers’ feedback or were disappointed after receiving misleading feedback. In addition, students’ attitudes towards feedback seemed to be different when the writing task was assessed.

8.2 Summary of Findings from Student Interview Results

The interview results suggest that the majority of the students enjoyed writing in their L1 and in their L2. As such, it appears that they have positive attitudes towards writing in general. Given that many reported writing in the L2 was
difficult compared to writing in the L1, writing in Japanese gave them a sense of self-achievement.

Despite the difficulties that writing in Japanese posed to these students, most perceived that they improved their writing over one semester and this was even more pronounced for those who continued on to study Japanese 4. The most frequently reported areas of improvement were grammar and *kanji*, with some reporting improvements in vocabulary. Students identified practice opportunities and feedback as helping them to develop their writing. They felt that individual assignments, writing the script for the oral assessments, weekly *kanji* tests and weekly activities from the textbook were particularly useful as they gave them opportunities to practice their writing. On the other hand, group assignments written in English and the very short weekly writing tasks did not contribute to their writing progress, as they did not allow sufficient writing practice. Japanese resources allowed students to obtain examples and learn new grammar and vocabulary; nevertheless, for some students the resources were too difficult. Peer feedback on group assignments, oral assessments and drafts of individual assignments were viewed as beneficial for writing development because students could notice their errors and improve their writing. However, peer feedback on the weekly writing task was not perceived as helpful and students felt they did not receive enough feedback from their teachers. In addition, it appeared that assessing peer feedback did not make a difference to the students’ views about peer feedback. The factors that influenced students’ writing development are summarised in Table 8.1.
Table 8.1

Factors that Influenced Students’ Writing Development

<table>
<thead>
<tr>
<th>Practice in writing</th>
<th>Beneficial</th>
<th>Non-beneficial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual assignment</td>
<td>Group assignment (in English)</td>
<td></td>
</tr>
<tr>
<td>Oral assessment</td>
<td>Weekly writing task</td>
<td></td>
</tr>
<tr>
<td>Weekly kanji test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weekly activities from textbook</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Reading in Japanese

<table>
<thead>
<tr>
<th>Reading in Japanese</th>
<th>Beneficial</th>
<th>Non-beneficial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provided examples (e.g., grammar pattern and sentence structure)</td>
<td>Too difficult, especially because too many kanji</td>
<td></td>
</tr>
<tr>
<td>Learned new vocabulary</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Feedback

<table>
<thead>
<tr>
<th>Feedback</th>
<th>Beneficial</th>
<th>Non-beneficial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher feedback:</td>
<td>Limited teacher feedback</td>
<td></td>
</tr>
<tr>
<td>When received feedback they were useful</td>
<td>Non-assessed peer feedback on weekly writing task</td>
<td></td>
</tr>
<tr>
<td>Non-assessed peer feedback:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group assignment (in Japanese)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oral assessment (writing a script)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assessed peer feedback</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Draft of individual assignment</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This section presented the responses given by participants in Japanese 3 (N = 40) and Japanese 4 (N = 22) in the semi-structured interviews. The next section will present the responses given by two teachers in Japanese 3 and one teacher in Japanese 4.

8.3 Teacher Interviews

Interviews with teachers were also conducted to determine their perspectives on writing development. One aim of conducting interviews, especially with the coordinator, was to determine how much emphasis was placed on writing at the lower-intermediate level. Another aim was to determine whether students’ perceived
improvement and factors that influenced their development corroborated with the teachers’ perspectives.

In Japanese 3 (Semester 1, 2014, the year of data collection), two teachers were interviewed, Hayashi-sensei (the coordinator) and Tanaka-sensei. In Japanese 4 (Semester 2, 2014), it was only possible to interview Hayashi-sensei as the other teacher had other commitments and was unavailable. In both Japanese 3 and Japanese 4, the teacher interviews were conducted after all student interviews were completed. A follow-up interview with Hayashi-sensei was conducted in early 2016, after all data had been analysed. All interviews were conducted in English. Hayashi-sensei was the coordinator of the course and therefore had control over the syllabus. The other teacher, Tanaka-sensei, was a sessional tutor (who was interviewed only in Japanese 3), and therefore did not have any control over the syllabus.

Analyses of the interviews are grouped into four broad themes: overview of the course design, teachers’ view on students’ writing development, factors that influenced progress in writing, and student background. All quotations presented are direct quotations from the participants, using pseudonyms.

8.3.1 Overview of the course design

Hayashi-sensei, the coordinator of Japanese 3 and 4, was asked about the course design in a follow-up interview in 2016. After analysing both the quantitative and qualitative data, it was important to obtain further information on course design to understand why certain assessment tasks and classroom activities were designed and allocated in a certain way and whether they were related to writing.

Hayashi-sensei designed the classroom activities and assessment tasks for Japanese 3 (Semester 1) and Japanese 4 (Semester 2). It is important to note here that in 2012, there was a major curriculum reform across the university that required all
language programs to teach language and culture together. The Japanese teaching team (i.e., all the coordinators in the Japanese Department) did not just implement this directive, but also decided to teach research skills. As Hayashi-sensei explained:

*In 2012, the focus of the re-structuring was to teach language and culture at the same time and we sort of decided that we need to teach some kind of research skills as well... so three things combined in one.*

This meant that Hayashi-sensei had to include tasks that supported students in learning research skills—he explained that these research skills were incorporated in the group assessment tasks.

It is interesting that in his interview in 2016, Hayashi-sensei reflected that his focus on the three areas may not have been allocated evenly:

*From 2012 to 2013, maybe to 2014, maybe I was focusing too much on the cultural side and maybe not enough language so from 2015 [the year after the research] I started to put focus back on to the language production as well.*

The research component and the writing task were linked; students were to conduct research as a group (using Japanese material) and complete the individual writing task on the topic of their research. As Hayashi-sensei commented:

*Doing research reading in Japanese... then from the research finding I intended students to write about that so that was a 15% writing task.*

The individual assignment for Japanese 3 was to produce a webpage. The reason for this task was that the coordinator believed it would help students with relatively low Japanese proficiency to develop their writing skills and produce something meaningful. He felt that adding pictures and hyperlinks could supplement some of the things that students could not express in their writing yet:
My intention was with limited ability to use Japanese they were able to come up with something meaningful for other people to read not just by writing a text. Pictures can basically supplement what they can’t precisely express in Japanese yet.

The individual assignment in Japanese 4 was a research-based assignment. Hayashi-sensei hoped that after completing Japanese 3 and Japanese 4, students would be able to write an academic essay in Japanese, with the assessment tasks in Japanese 3 providing the relevant practice:

The goal for both Japanese 3 and Japanese 4 combined was to be able to write an academic essay. Web design was just a practice... yes I designed like practice. In Japanese 3 I don’t think students have enough language to write a coherent piece of work. So that is why the webpage can be just a compilation of different pieces of information whereas an academic essay has to be a coherent piece of work so yes, Japanese 3 was designed to be a practice.

Hayashi-sensei’s intention for the group assignments was for students to undertake the research together, which would allow them to collect more information than they would have if each student researched individually. He hoped that by having more resources, students would then be able to write a better paper when writing their individual assignment:

For one person to come up with a set task is quite difficult so what I was trying to do was to get three people together... they can all research different aspects and by combining all the resources each individual can then come up with a better quality assignment. So my purpose was to get students to do more research and get a combined effort together.
Japanese 3 included three group assignments—out of the three, two were completed in English. Similarly, in Japanese 4, students were to write both group assignments in English. When Hayashi-sensei was asked why these assignments were written in English, he explained that they were comprehension exercises rather than writing exercises: ‘the reason was simple, we wanted to make sure students understood what they read so having them summarise the findings in English was the easiest way’.

Other than the individual assignments and group assignments, each subject included two oral assessments. One of the speaking assessments in each subject was to make a group video. Although the assessment itself was on the speaking component, students had to write a script before making the video (the script was not assessed). Hayashi-sensei hoped that writing a script before making the video for the oral assessment would help students to prepare correct sentences and memorise them for the role play, ‘role play is a prepared speech. Because students have to memorise the lines, we would like them to memorise the correct sentences’.

The classroom activities in both Japanese 3 and Japanese 4 consisted of weekly reflective writing tasks (three sentences) and weekly kanji tests. Both the reflective writing tasks and kanji tests were not assessed. Kanji tests were dictated; this included the KCs that students had to learn each week—the teacher read out the whole sentence and students had to write it down.

Hayashi-sensei assigned students to do the reflective writing tasks in both subjects because he felt this would be beneficial for improving students’ writing skills. In Japanese 3, weekly reflective tasks were seen as providing students with practice in writing something ‘on the spot’, which they had never done before:
Students had a chance to write a paragraph in class on what they have learnt in that week. For many students that was the first time that they had to actually write something on the spot.

Hayashi-sensei did not assign a specific topic for this reflective writing activity, as he wanted students to reflect on what they learned in class and then write about it:

The original idea came from the education theory on self-reflection... the purpose was to review what they have learnt through the week and actually remember trying to put that, commit that into their memory. So that is why I didn’t set a specific task. Think about what they have done in the class then whatever they thought was most interesting they can just write about it.

8.3.2 Teachers’ views on L2 Japanese writing

Teaching writing at an early stage is considered important for character-based languages such as Japanese and Chinese (Cook & Bassetti, 2005). Both Hayashi-sensei and Tanaka-sensei, when asked whether they thought that the teaching of Japanese writing was important at the lower-intermediate level, concurred that it was. Hayashi-sensei emphasised that because there are many writing assignments in Japanese 5, it is necessary for students to gain writing practice at the Japanese 3 and Japanese 4 level. He explained:

I have to say it is important because after Japanese 4 as soon as students go into Japanese 5 intermediate level, they are asked to write a whole essay in Japanese. Not just essays but different styles of writing. So if I don’t teach writing skills in Japanese 3 and 4, they will be totally lost in Japanese 5.

As Hayashi-sensei was the coordinator, his response was more curriculum-focused and concerned about preparing students for the next step in terms of their
learning progression within the institution (i.e., writing assessments in Japanese 5).

In contrast, Tanaka-sensei’s response related to language learning in general, emphasising that it was important for students to learn all language skills, including writing:

*I believe all language learners benefit from learning that integrate the four skills (speaking, listening, writing and reading) suitable to their competency level. Also, writing is important academic skill and one that is used for majority of assessment.*

In the follow-up interview in 2016, when discussing the course design, Hayashi-sensei also commented on the importance of *kanji*, grammar and the ability to structure an essay in L2 Japanese writing at the lower-intermediate level. Hayashi-sensei said:

*By the end of Japanese 4, there is a clear difference between students who can use kanji sufficiently and not. So the first hurdle is to see how many kanji characters students can read or write.*

Nevertheless, Hayashi-sensei admitted that it was difficult to teach *kanji* and he was unsure about the best way to assist students learn *kanji* effectively. He noted that ‘some students seem to think that I have to explain to them, how *kanji* is made, how to memorise so I am confused too’. He also admitted that ‘I didn’t really teach them. I just told them you have to know this page’.

Being able to apply grammar is also important for improving writing. Hayashi-sensei said:

*Then the second task is grammar [what the teachers have to look at in terms of L2 Japanese]. How to put grammar, how students can put different words in correct order to make sense.*
Although kanji and grammar are important, Hayashi-sensei explained that he focused more on teaching students how to structure an essay, rather than focusing on kanji or grammar:

*Then there is another level. How to put a research essay together. You have to have a structure. So I think I have put more emphasis on writing a structured essay talking about the structures of writing but perhaps not enough emphasis on kanji or grammar.*

### 8.3.3 Teachers’ views on students’ writing development

When interviewed at the end of Japanese 3, Hayashi-sensei believed that the students who completed Japanese 3 showed kanji development during the semester. His beliefs related not only to this group of students, but also to past groups, given that he had been teaching Japanese 3 for several years. However, he also believed that among the Japanese 3 students, there was not much development in other areas of writing, such as in writing sentences. He felt that at the lower-intermediate level, students seemed to be more focused on learning kanji rather than developing fluency in writing:

*I think in Japanese 3 most of the writing development comes in terms of learning more kanji and not necessarily actually writing sentences. Most of the students are in the mindset that Japanese writing is mostly done in kanji so learn more kanji than develop their fluency and in terms of writing sentences.*

He believed that students were still unable to write well in Japanese 3 and noted that ‘they haven’t really learned so much. I don’t think they can write or express themselves fluently’. He continued, ‘I think they are still working on their
speaking skills in Japanese 3’, meaning that at this level, most students try to develop their speaking skills rather than their writing skills.

Tanaka-sensei commented that she was not able to tell whether the students’ writing developed over the semester due to her limited interaction with the students. She admitted:

*I am not sure to be honest. Because I really hadn’t had a chance to see their level at the beginning and I hadn’t had the time to individually assist students.*

It is important to note that this was Tanaka-sensei’s first year of teaching at this university and she only taught one class in Japanese 3. This did not give her much exposure to students’ performance and their development.

In Japanese 4, Hayashi-sensei was asked about students’ overall areas of improvement in writing (e.g., *kanji, hiragana, structure, content*) after completing Japanese 3 and 4. He said that rather than considering their writing development by examining individual aspects of writing ability, he saw their development more holistically, in terms of whether they could express what they wanted to communicate in different ways:

*I never thought of that in terms of writing ability because what happens in Japanese 3 and 4 is students get to learn more complex grammar patterns so instead of just saying, like ‘my name is such and such’... most of the development I see is whether students can and has begun to express different views, rather than writing ability.*

**8.3.4 Factors that influence progress in writing**

The teachers were also asked for their insights on what they considered could lead to progress in writing. The teachers mentioned several factors, some of which were similar to those mentioned by students: practice, reading, motivation and
feedback (both teacher and peer feedback). Each factor will be discussed in terms of whether or not the teachers felt that it could explain their students’ writing improvement, and on reflection, some of the students’ challenges.

8.3.4.1 Practice in writing

Both teachers noted the importance of providing students with opportunities to practice for them to improve their Japanese writing. Hayashi-sensei commented that development depended on the amount of practice: ‘with any of the language studies, how well students develop is very much dependent on how much practice they have. So I would like to give more writing practice’. Tanaka-sensei also believed that practice was essential to improve writing. She thought that providing good examples was important:

*I think they need a good model [sample] in various genres. Then, the text should be analysed either in class or groups and followed by practise, editing.*

In terms of the assessment tasks for Japanese 3, Hayashi-sensei believed that individual assignments provided the most help to students for improving their writing. This was because they were designed specifically to develop writing skills. Nevertheless, he did not know if they had been as effective as he anticipated. Hayashi-sensei thought that some students did not do much research before writing up their essays:

*For those students [who did not do research] obviously it was not useful… research task was supposed to be an input task so if students didn’t do the input from the research there is no effect.*

However, Hayashi-sensei also thought that students faced difficulties because the resources they used might have contained vocabularies and grammar they had not been exposed to yet in Japanese 3:
Perhaps dealing with unknown vocabulary, they had to find out different words and meaning of different words and actually using them without really knowing how to use them. Also sometimes the grammar, things like passive[form] is not taught in Japanese 3 but because many websites use that and some students tried to emulate that.

Tanaka-sensei believed that many students struggled with structuring their webpage (i.e., the Japanese 3 individual assignment). This may be because they did not have a sample to guide them. In addition, it seems that many students were unsure about their audience was and thus they encountered difficulty in their writing. Tanaka-sensei also noted that the websites the students referred to were authentic Japanese websites and were not particularly suitable for their assignment:

Some students got the idea of the target audience but I think a lot of them didn’t... I think it is because there wasn’t any model or sample for them so it was either their own work or authentic Japanese site and there wasn’t anything in between.

In terms of the group assignments in Japanese 3, Hayashi-sensei thought that these did not really help to improve students’ writing skills. The group assignments were designed to develop students’ research skills (see Section 8.3.1) and were not focused on developing their writing skills. Hayashi-sensei commented:

I am sure there were students or some really focused groups who produced their writing and have it read by their other group member... so that they were conscious about improving their writing skills... but in many cases, they didn’t have much effect on their writing skills—most of the writings on the group tasks were presented in English—yes phase 2 and 3, just the phase 1 was in Japanese so.
Tanaka-sensei believed that if a group worked well together, students would be more likely to improve their writing structure. By being able to see how other students structured their writing, students could learn from one another: ‘I think structure-wise definitely because they can copy from each other I suppose. And maybe vocabulary as well’.

Teachers were also asked whether writing the script for the oral assessment helped students develop their writing skills. Hayashi-sensei said that it might have been beneficial for many students, as they had to write their own script. Moreover, if students had the opportunity to have their writing proofread by someone else, this may have helped them to notice some errors in their writing and improve their writing. He commented:

*Writing a script was a helpful part that many students say, I am sure they can come up with their own script and probably have it proofread by someone... so they have picked up something.*

Interestingly, Tanaka-sensei thought that whether students benefited from the script writing exercise depended on the group they were in. She believed that those students in groups who worked hard in writing the script did benefit. She said:

*I think it depends on the group, some group put a lot of effort into it. And some kind of did it on the spot, so it depends with the group whether it helped with their writing or not.*

In terms of the classroom activities, Hayashi-sensei thought that the reflective writing task was not as effective as he thought it would be. He admitted that it may have been difficult for some students to write as many may not have previously been asked to write without prior preparation. Further, even the students who were able to
reflect on what they had learned may still have found it difficult to put their reflections into words:

For many students that was a first time that they had to actually write something on the spot. In Japanese 1 and 2, they only do an assessment piece yes so they have a time to think, write, having proof read and then submit for assessment but for Japanese 3, especially for in-class writing that was the first one, so some students find it quite difficult to think what they have learnt and what they have done that week and actually link that to the words.

Tanaka-sensei believed that at the lower-intermediate level, this kind of activity was still very difficult for students to accomplish, and this was evident in the classrooms observed by the researcher. When students were undertaking this activity, they seemed unsure what they should be writing about—some students took a long time to write something, and others sat there thinking about what to write. In contrast, students seemed able to write something promptly when given a question or a topic, as was evident when answering the workbook questions in class. Tanaka-sensei commented:

I know what the intention is to for the students to think and integrate what they have learnt in lecture, Seminar 1 and Seminar 2. And then write up something but I don’t think they are used to that kind of activity for some of the students especially from the ones from high school it was difficult. They couldn’t think of anything to write.

As many students were unfamiliar with this kind of writing task, Tanaka-sensei thought that perhaps giving them a specific task would have been better:

I think at this level for some students it may have been easier if they were given a specific task rather than reflect on what they have learnt.
Hayashi-sensei’s view on the reflective writing task seemed to change after Japanese 4 (Semester 2, 2014). He said, ‘I used to think that it is OK, the three sentences, but after doing that for a year I don’t think it is enough’.

In the follow-up interview in early 2016, Hayashi-sensei was asked whether he still used the reflective writing task in his Japanese classes. He explained that he changed the reflective task from writing to speaking, as he was not satisfied with the writing task in 2014. However, he still struggled to find a task that would enable students to develop their writing skills, ‘writing, yes, I changed the in-class assessments slightly… but from the final exam results I don’t think it was very successful’.

Hayashi-sensei was yet to decide whether to provide a different writing exercise in 2016, mainly because he had difficulty in allocating extra time for writing. He stated, ‘I really don’t know what else to do. I can probably give more writing in class but as it is quite pressed for time already, I don’t know if I can change anything’.

In Japanese 4, Hayashi-sensei was also asked about his opinion on the weekly kanji test. He believed that the weekly kanji test helped students improve their writing, but as students had to learn many KCs, he was uncertain how much students were able to retain. He commented, ‘this year we learnt 200 kanji all together—100 in Japanese 3 and 100 in Japanese 4. How many of those 200 kanji students remember is the question’. Nonetheless, he felt that ‘kanji is important… even if they cannot write, at least the reading ability is necessary’.
8.3.4.2 Reading in Japanese

Another factor that both teachers perceived as important for students to develop their writing was reading. Hayashi-sensei noted the importance of input in language learning. To produce something, appropriate input is vital:

Language learning is about getting the input first and then processing that information and producing at the end. Research task was supposed to be an input task so if students didn’t do the input from the research there is no effect.

In Japanese 3, he commented that students should read more before writing in order to start writing well in Japanese 4. This is illustrated in the comments below:

I put more emphasis on readings than anything else in Japanese 3 believing that in order to write something to produce something, students must have an input and I see reading as the type of input. Hopefully they will do more writing or output work in Japanese 4.

Tanaka-sensei mentioned that it was important to first look at a particular genre to have an understanding of the way it is written before actually start writing:

My personal opinion is to teach writing, you have to teach from reading. You have to spend time analysing a particular sample from a particular genre and they go over it together and then you can move onto writing exercises.

In order to write up Stage 1 of the group assignment in Japanese 3, students needed to read Japanese websites. Hayashi-sensei believed that this helped some, but not all, students with their kanji learning. He also hoped that by providing reading opportunities, students were able to learn how Japanese sentences were structured:

It depends on students, some students are very conscientious, they have really applied themselves to try to read something in Japanese. Obviously for those
students they have acquired more kanji and hopefully they became aware of how different bits are tied together in Japanese sentences or within a text or paragraph.

Tanaka-sensei thought that reading authentic Japanese websites would have been difficult; therefore, at this level students would not obtain much help from these websites:

For the students I know reading authentic Japanese sites was a bit difficult especially with kanji, if they could find a site that is targeted to their level, it would have helped more.

Hayashi-sensei also thought that the Japanese websites might have been difficult for some students to understand:

Probably yes [i.e., that the website was too difficult] my intention was for them to find something that they can read by themselves but obviously for students who don’t have enough time to look for something that they can access, they just pick anything. They tried to read it whatever reads so that didn’t quite work so I need to think the assessment for next year.

In the follow-up interview, Hayashi-sensei commented that although he believed that input (i.e., reading) was important for writing development, the reading assessment in 2014 (i.e., reading resources in Japanese and writing in English) was not very successful. Hence, instead of having a reading assessment (reading in Japanese and summarising in English) as part of the group assessment, he incorporated a peer evaluation\(^3\) in 2015 (the year after the research was conducted).

\(^3\) ‘Peer evaluation’ differs from ‘peer feedback’. In peer evaluation, students evaluate their peers on their contribution to a group assignment, so that everyone is satisfied: this was part of the assessments conducted in 2015.
8.3.4.3 Motivation

Empirical studies (e.g., Cumming, 2012; Sasaki, 2011) suggest that motivation affects students’ writing. Motivated learners tend to improve more than those who are less motivated (Cumming, 2012). Both Hayashi-sensei and Tanaka-sensei also identified that motivating students to write was important, especially for learners at the lower-intermediate level. As Tanaka-sensei commented of motivation, ‘in my experience, it seems students have love and hate relationship with writing. So it may be important to motivate students who do not like writing’.

When asked what kind of tasks would motivate students to write, Hayashi-sensei commented that activities such as discussion boards do motivate students, as they can receive feedback on what they have written:

*Activities that students that can get a reply—like discussion board or in the old days like receiving a letter, actually getting replies from Japanese students. Nowadays it has shifted to discussion board.*

Nevertheless, motivating students and keeping them motivated can be a difficult task: ‘that [e.g. discussion board] motivates them at the start but actually maintaining the motivation is difficult’.

Similarly, Tanaka-sensei also thought students were motivated if they could practice something related to a real life situation. Tanaka-sensei cited one of the Japanese language courses at another university as a good example of successful student motivation. A project-based approach with authentic writing and speaking tasks seemed to motivate students:

*The ones that are close to real life motivates more. I am thinking of the capstone course at the University of NSW where they have to give presentation to the Japanese audience and they have to send invitation in*
Japanese, (so they have to write in Japanese) and then presentation materials have to be in Japanese and handouts to the audience and also thank you letter to the audience that really motivates the students.

Both Hayashi-sensei and Tanaka-sensei’s comments indicate the importance of providing an interesting task or topic to motivate students to write.

8.3.4.4 Teacher feedback on writing

Teacher feedback was perceived as important by both teachers for students’ writing development, yet they acknowledged that they faced the challenge of ‘restricted time’, which limited their ability to provide adequate feedback.

For example, Hayashi-sensei believed that providing feedback was useful, ‘feedback is useful. It is a necessary part of improving someone’s writing’. But he admitted that his time was restricted and this meant that he provided limited feedback. He felt that he should provide more feedback to students, not just commenting on grammatical errors:

The time is quite limited to give feedback... I know it will be good to sit down with individual students but we can’t do this, so I usually limit myself just giving feedback on grammatical errors like using different particles and tense but I know I should do more.

He also admitted that while providing feedback on kanji and grammar was not difficult, providing feedback on written work (e.g., short essays) could become very difficult due to the number of students he taught:

Writing is very hard to check if I give them as a homework. Kanji test is fine I can check the test. Grammar is the same too, I can give them grammar exercise to check but with writing I have to actually read it. With 240 students that is just not possible.
Tanaka-sensei also explained the importance of feedback—it allows students to confirm whether they are writing appropriately or not. She commented, ‘I think it is important for them to know whether they are writing grammatically correct sentences or not’. However, Tanaka-sensei also stressed that she had limited time and was unable to give enough feedback. Tanaka-sensei explained that when providing feedback for the reflective writing task in class, instead of giving detailed feedback, she could only indicate areas that were correct or incorrect by using different coloured highlighters. Tanaka-sensei suggested that discussing common mistakes and providing examples in class would have been a better alternative:

*With the allocated time the only thing I could do was to highlight. And also encourage them to write so that they know they have done it good but marking it in green was the only way I could do it within the given time and I would definitely like more time to give them feedback... individually is difficult but if we can focus on common mistakes or if I could have shown good examples to the group, that would have been good.*

She also thought that a peer group working well would assist in students’ writing development; she believed that students would try to help each other. On some occasions, she believed that students would listen to their peers more than their teacher:

*I is always good if a group works well. It is always helpful to get scaffolding from peers. Sometimes they are more likely to listen to their peers than teachers.*

Hayashi-sensei also commented on the usefulness of peer feedback, and this explains why he incorporated peer feedback as one of the assessment tasks in Japanese 4.
8.3.4.5 Peer feedback on writing

In Japanese 4, each student had to give peer feedback on two drafts of individual assignments (the peer feedback represented 5% of their total mark). Although instructions on how to give feedback were provided during lectures, they were very brief and students did not get any practice during the semester before undertaking this task. Students had six compulsory criteria for the peer review task that required answers; most were broad questions with a very limited language focus, as shown in Table 8.2.

Table 8.2

Peer Feedback Criteria

1. At which point did you feel most interested by this piece? When least? Explain.

2. Do you feel this paper relies on evidence, or on opinion or intuition? If the latter, cite examples of where this paper relies on opinion and intuition and give suggestions as to how the writer can write more objectively.

3. Does the paper answer the questions set out in the assessment task?
   i. General information/background to your chosen area of Japanese writing system;
   ii. How the chosen aspect of the writing system developed over the years;
   iii. How it compares to writing systems in other cultures/languages; and,
   iv. Why it is important for learners of Japanese to know this information, or how it affected your learning of Japanese language.

4. Can the conclusion of this paper be convincingly drawn from the thesis and the argument made in the body of the paper? Why or why not?

5. Overall comments (e.g., grammatical errors, structural errors and/or good points about this essay)

6. How would you rate this essay?
Hayashi-sensei believed that peer feedback was useful as students could obtain different kinds of feedback. He thought this was good preparation for real life situations. Further, he stressed the importance of having an audience in mind when writing. He thought the peer feedback exercise would assist students to develop this kind of mindset, as their writing will have been read by people other than their teachers:

*Peer feedback is a useful assessment task for students to get familiar with the assessment tasks... What usually happens at a university is when students write an assessment task they only have their teacher to think about but in real life it’s slightly different... like when you and I write an essay or an academic journal article, we have to think about a particular type of audience. So peer review is partly a preparation for... to resemble something like that. Get someone in the same class to read your writing.*

Hayashi-sensei explained that that peer feedback not only helped students improve their writing but gave them reading practice and exposure to different ideas:

*Also from the reading side, obviously it is reading practice, reading comprehension practice, by reading other students’ tasks they can also get different point of views... By reading other people’s papers and see how they make their research more relevant to their reader they can also get the ideas on how to better their own essays. Suppose to be a win-win strategy.*

He further commented that peer feedback worked better in improving writing content and structure than grammar:

*Peer feedback works quite well with the contents but not so well with grammar. Work with some students but not with many students... I think*
students were able to correct each other’s structure and contents rather than grammar.

Hayashi-sensei believed that giving feedback itself would not have been difficult for the students, but it is questionable whether the feedback was useful or not: ‘for students to get marks for this peer feedback task they only had to write a certain amount of words, I didn’t really question their quality’. This means that as long as the students provided feedback, they were given acceptable marks for this assessment. When asked why the six compulsory questions students had to answer mostly related to content (see Table 8.2), Hayashi-sensei responded that the guidelines used for peer review were adapted from guidelines used by non-language subjects, such as multimedia and communication, at the university. The feedback that students provided for those subjects was focused on the essay’s content and structure. Further, the compulsory criteria listed in Table 8.2 were based on guidelines provided by a computer program called ‘Turnitin’ on how to establish a peer review task. Suggestions are given on what kind of questions can be included, and these were mainly content-related. He also believed it would have been very difficult to prepare grammar-based questions. As Hayashi-sensei explained, ‘I did put questions saying, “Is there any grammatical errors”… but I can’t really write a question saying, “Is there a sentence pattern saying using such and such”’.

He thought that not every student benefited from the peer review task. Some students did not read the comments made by their peers:

I don’t think that 50% of the students did much after receiving the comments.

There is a function that I can see whether the students had actually read the comments but not everyone comes out as ‘read’.
Hayashi-sensei was unsure whether the students wrote better texts after they had received feedback from their peers:

_I don’t know if they... most of the students made revisions after they received the comments but whether that was for better or worse I don’t know. I really haven’t done ... looked into that much._

### 8.3.4.6 Student background

Understanding the differences between learners from various backgrounds is important in countries like Australia, and it was clear from the interview responses that both teachers were aware of that. They noted how they were conscious that students originating from character-based backgrounds could have an advantage over those from non-character-based backgrounds particularly in terms of learning kanji. However, they argued that sometimes acquiring kanji depended on the student. As Hayashi-sensei noted:

_Their backgrounds do affect their kanji ability. It’s really hard to generalise, there are some non-kanji background students who are good at kanji, obviously they know how to read and write and how to form a kanji and there are some Chinese-background student who don’t want to learn Japanese writing._

Tanaka-sensei observed that although many characters in Chinese and Japanese were the same, many also appeared similar but were not the same. This led some character-based background students to make incorrect assumptions due to these similarities, and use Chinese characters that did not exist in Japanese. Thus, any advantage background learners may have had initially might eventually become a disadvantage. As she said:
I think initially the ones with kanji background find it easy to do the writing task but then some of those students, because in the beginning it was easy for them, they become a bit lazy and they keep using Chinese characters. They always get marks deducted for that.

Hayashi-sensei thought that everyone was capable of developing their kanji skills and ultimately this would depend on each individual, rather than their background:

*It’s more to do with the personality rather than the background of students.

There are students’ who just finished high school and suddenly being burdened with so many kanji every week... some students find Japanese 3 relaxing. In terms of learning kanji and writing kanji, yes it has to do with each individual and their own experiences of learning Japanese elsewhere I think.*

### 8.4 Summary of Student and Teacher Interview Results

One aim of conducting interviews with both students and teachers was to determine whether their views on writing development aligned. In addition, it was also important to investigate whether students’ perceptions of writing development corresponded with the focus of the writing (assessment tasks and classroom activities) for Japanese 3 and Japanese 4 subjects. In the above sections, the results from student interviews and teacher interviews were provided. In this section, students and teachers’ views will be compared and discussed.

Out of the assessment tasks in Japanese 3 and Japanese 4, individual assignments appeared to be the task that Hayashi-sensei designed to improve students’ writing skills in particular. The majority of students thought that these assignments were useful in improving their writing (see Section 8.1.3.1). While
Hayashi-sensei thought that individual assignments assisted in improving students’ writing, he thought they varied in their effectiveness. He believed that these assignments led to improvements in the writing of those students who conducted their research properly before writing their individual assignment.

Other assessment tasks were not designed specifically to improve writing skills. For example, the focus of the group assignments was more on developing students’ research skill. Students reported that their writing skills did not improve from completing these assignments. However, most students reported that oral assessments in both subjects were beneficial for their writing, as they had to create a writing script.

One of the classroom activities that Hayashi-sensei incorporated to develop students’ writing skills was the weekly reflective writing task. However, he admitted that it was not as effective as he had anticipated, and this was reflected in students’ comments. Students believed that the task, consisting of only three sentences, did not provide enough practice. Further, Tanaka-sensei believed that the spontaneous and ‘freeform’ nature of the task was difficult for students to accomplish; many had trouble initiating ideas. This was corroborated by the researcher’s observation notes.

Three main influential factors for writing development were reported by both the students and teachers: practice, feedback and reading. The interview results reveal that students and teachers’ views on these factors aligned. That is, practice is critical for improving writing. In addition, feedback, especially teacher feedback, is important, as it enables students to notice their linguistic gaps. In the current study, both students and teachers felt that insufficient teacher feedback was given. Finally, appropriate input through reading also allowed learners to develop their writing skills. The interview results suggest that reading was useful in improving students’
writing, as the texts provided examples. However, on some occasions when students were unable to comprehend a text (e.g., complex website text in Japanese), they were ineffective. In addition to the three main factors discussed here, the teachers also identified motivation as a critical factor for writing development. If students were provided with interesting writing activities, it is likely that they would become motivated and thus engaged in the task.
Chapter 9: Discussion of Study 2 and of the Entire Research Project

The classroom-based longitudinal study (Study 2) investigated the writing development of L2 tertiary Japanese learners over one semester, as well as over one year. The study also explored the perspectives of participants (both the students and teachers) on writing development.

The first section of this chapter discusses and synthesises the main findings of Study 2 in light of the two RQs. In order to answer the two RQs, six different aspects of writing were examined: kanji complexity, syntactic complexity, accuracy, fluency, content and structure. For the first RQ, the discussion mainly centres upon whether development was evident in students’ writing over the short and long term (i.e., over one semester and one year). This is followed by a discussion on whether the findings from the text-based analysis align with students’ perceived improvement and teachers’ perceptions. Regarding the second RQ, the difference in writing development between character-based learners and non-character-based learners is discussed.

The second section of this chapter provides an overall discussion of the project. The results of Study 1 and Study 2 are compared and discussed.

9.1 Research Question 1:

Does writing ability among university students improve in the short term (one semester), as well as in the long term (one year). If so, which aspects of writing develop? Does a study of longer duration make a significant difference?

In order to investigate whether lower level intermediate students improved in their writing over the short term (i.e., one semester), the study examined the writing
of students who completed Japanese 3 (Semester 1, \(N = 41\)). To examine the long-term writing development (i.e., one year) the study examined the writing of students who completed both Japanese 3 and Japanese 4 (Semester 1 and Semester 2, \(N = 22\)). The students’ performance on pre- and post-tests were used to judge improvements in the short and long term. As explained in Chapter 6 (see Section 6.2), the pre-test in Japanese 3 was held at the beginning of the semester (Week 1, Semester 1) and the post-test was held at the end of the semester (Week 12, Semester 1). Those students who continued to Japanese 4 were asked to do the same writing task at the end of Japanese 4 (Week 12, Semester 2), and this writing served as the post-test for the one year study. The writing task topic, ‘Write a promotional newspaper or magazine article about your hometown or a place you know very well’ was given for both tests.

Various measures were explored in relation to students’ writing development. This section offers a discussion based on the quantitative findings of kanji complexity, syntactic complexity, fluency, accuracy, content and structure for short- and long-term results. Tables 9.1 and 9.2 summarise the findings, presenting the measures that showed improvement, no change or deterioration. The results are discussed in the following sections. Where differences existed between the short- and long-term results, they are compared and discussed. Participants’ perspectives (both the students and teachers) on writing development are also provided throughout the discussion; these may explain why some writing features improved, while others did not.
Table 9.1

*Short-term Results (One Semester)*

<table>
<thead>
<tr>
<th>Measures that showed improvement</th>
<th>Measures that showed no change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measures</td>
<td>Effect size</td>
</tr>
<tr>
<td>1) Kanji complexity</td>
<td></td>
</tr>
<tr>
<td>KC RTTR</td>
<td>large</td>
</tr>
<tr>
<td>KW RTTR</td>
<td>large</td>
</tr>
<tr>
<td>KC CTTR</td>
<td>large</td>
</tr>
<tr>
<td>KW CTTR</td>
<td>large</td>
</tr>
<tr>
<td>2) Syntactic complexity</td>
<td></td>
</tr>
<tr>
<td>Ratio of te-form</td>
<td>medium</td>
</tr>
<tr>
<td>Ratio of complex sentences</td>
<td>large</td>
</tr>
<tr>
<td>Ratio of clause conjunction</td>
<td>large</td>
</tr>
<tr>
<td>Ratio of propositional endings</td>
<td></td>
</tr>
<tr>
<td>Ratio of epistemic endings</td>
<td></td>
</tr>
<tr>
<td>Ratio of expressive endings</td>
<td></td>
</tr>
<tr>
<td>3) Accuracy</td>
<td></td>
</tr>
<tr>
<td>Ratio of EFCs</td>
<td>large</td>
</tr>
<tr>
<td>Words per error-free sentences</td>
<td>medium</td>
</tr>
<tr>
<td>Ratio of error-free sentences</td>
<td></td>
</tr>
<tr>
<td>4) Fluency</td>
<td></td>
</tr>
<tr>
<td>Number of words</td>
<td>large</td>
</tr>
<tr>
<td>Number of sentences</td>
<td>large</td>
</tr>
<tr>
<td>Number of clauses</td>
<td>large</td>
</tr>
<tr>
<td>Number of KCs</td>
<td>large</td>
</tr>
<tr>
<td>Number KWs</td>
<td>large</td>
</tr>
<tr>
<td>6) Content</td>
<td></td>
</tr>
<tr>
<td>Sufficiency</td>
<td>large</td>
</tr>
<tr>
<td>Persuasiveness</td>
<td>large</td>
</tr>
<tr>
<td>7) Structure</td>
<td></td>
</tr>
<tr>
<td>Coherence</td>
<td>large</td>
</tr>
<tr>
<td>Paragraphing</td>
<td>large</td>
</tr>
</tbody>
</table>
Table 9.2

*Long-term Results (One Year)*

<table>
<thead>
<tr>
<th>Measures that showed improvement</th>
<th>Measures that showed no change</th>
<th>Measures that showed deterioration</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Measures</strong></td>
<td><strong>Effect size</strong></td>
<td></td>
</tr>
<tr>
<td>1) <em>Kanji</em> complexity</td>
<td>KC RTTR</td>
<td>large</td>
</tr>
<tr>
<td></td>
<td>KW RTTR</td>
<td>large</td>
</tr>
<tr>
<td></td>
<td>KC CTTR</td>
<td>large</td>
</tr>
<tr>
<td></td>
<td>KW CTTR</td>
<td>large</td>
</tr>
<tr>
<td>2) Syntactic complexity</td>
<td>Ratio of complex sentences</td>
<td>large</td>
</tr>
<tr>
<td></td>
<td>Ratio of clause conjunction</td>
<td>large</td>
</tr>
<tr>
<td></td>
<td>Ratio of te-form</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ratio of propositional endings</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ratio of epistemic endings</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ratio of expressive endings</td>
<td></td>
</tr>
<tr>
<td>3) Accuracy</td>
<td>Ratio of EFCs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Words per error-free sentences</td>
<td></td>
</tr>
<tr>
<td>4) Fluency</td>
<td>Number of words</td>
<td>large</td>
</tr>
<tr>
<td></td>
<td>Number of sentences</td>
<td>large</td>
</tr>
<tr>
<td></td>
<td>Number of clauses</td>
<td>large</td>
</tr>
<tr>
<td></td>
<td>Number of KCs</td>
<td>large</td>
</tr>
<tr>
<td></td>
<td>Number KWs</td>
<td>large</td>
</tr>
<tr>
<td>6) Content</td>
<td>Sufficiency</td>
<td>large</td>
</tr>
<tr>
<td></td>
<td>Persuasiveness</td>
<td>large</td>
</tr>
<tr>
<td>7) Structure</td>
<td>Coherence</td>
<td>large</td>
</tr>
<tr>
<td></td>
<td>Paragraphing</td>
<td>large</td>
</tr>
</tbody>
</table>

**9.1.1 Kanji complexity**

The measures used to investigate development in *kanji* complexity were KC RTTR, CTTR and KW RTTR and CTTR. These measures are ratios and not frequency counts (see Chapter 3).

Findings for both the short and long term revealed that students’ writing on all measures of *kanji* complexity improved (see Tables 9.1 and 9.2) and the...
improvements were all statistically significant (p < .001; see Chapter 7, Section 7.1), with a large effect size. Further, for both KCs and KWs, the mean difference between the pre- and post-test was greater for the long term compared to the short term. For example, the mean difference for KC RTTR for the short term was 1.18, whereas it was 1.86 for the long term. This suggests that the improvement over the long term was greater than that in the short term. Improvement in kanji complexity is different from simply using more kanji in a text (i.e., fluency); it means that learners could use a wider range of KCs and KWs. In other words, the number of KCs and KWs includes different types.

The improvement in kanji complexity in both the short- and long-term results may be attributed to a number of activities in which students engaged: weekly kanji tests (non-assessed), ERs and group and individual assignments. According to Nation (2001), three important processes are necessary for learning and remembering new words: noticing, retrieval and generative use. Preparing for the kanji tests exposed students to new vocabulary (noticing) as did ER. Retrieval and generative processes were activated when students wrote assignments.

Both Japanese 3 and Japanese 4 included weekly kanji tests, and students were given approximately 8 to 10 kanji each week, which they were required to memorise before the test. These kanji were selected from textbook chapters the students studied each week (one chapter every week). In Japanese 3, students were introduced to 100 new KCs and in Japanese 4, they were also introduced to another 100 new KCs (these were the kanji on which students were tested). The weekly kanji test was a dictation element, and the teacher read out the sentences (including the kanji for that week). More than half of the students felt that their kanji improved and attributed this improvement to the weekly tests. From the teacher’s perspective, only
Hayashi-sensei commented about students’ writing development. He felt that students showed kanji development in both semesters. However, he admitted that teaching kanji was difficult and he was unsure about the best way to assist students to learn kanji beyond memorising the sets given for the weekly tests.

In Japanese 4, students also had dictation tests of kanji they had learned previously (i.e., in Japanese 3). This meant that the students’ knowledge of kanji was being revised. As such, one reason for students’ improvement being greater in the long term compared to the short term could be due to their exposure to more KCs after one year of study. According to Nation and Waring (1997), in order for a learner to insert new vocabulary into their long-term memory, they must encounter the word several times. By accumulating more KCs after one year, students were able to use a variety of kanji in their Japanese 4 post-test compared to the Japanese 3 post-test.

Although the kanji test was not assessed, students had to report on their kanji test mark in front of everyone immediately after the activity in class in both subjects. Perhaps this public reporting also encouraged students to memorise the required kanji every week. This is reflected in comments by Alex (Japanese 3), ‘sensei actually asked for the marks so probably that is why people put in effort’.

Another factor that can explain the improvement in kanji complexity is the ER that students were required to do in both the short (one semester) and long term (one year). Students were required to do weekly readings before each lecture (non-assessed) and frequent readings were required in classes, as well as readings in preparation for their assignments. A large number of studies on ER (e.g., Lee & Hsu, 2009; Mermelstein, 2015) indicate that ER leads to significant improvements in

---

4 The other sessional tutor in Japanese 3 was new and only taught one tutorial in Japanese 3. She could not comment on students’ development.
learners’ vocabularies. Repetition and the number of encounters are the primary factors in vocabulary learning (Nation, 2015). It is likely that ER enabled the students to encounter the *kanji* they had learned during the semester/s, and this assisted them in acquiring those *kanji*. Moreover, students would have been able to accumulate a wider variety of KCs in the long term than in the short term and therefore, were able to use more *kanji* over the long term (hence the greater improvement).

**9.1.2 Syntactic complexity**

A number of measures were used in this study to examine syntactic complexity. These were the ratio of te-form, the ratio of complex sentences, the ratio of clause conjunctions, the ratio of propositional endings, the ratio of epistemic endings and the ratio of expressive endings. The findings for syntactic complexity were not clear. Some measures showed improvement only in the short term (e.g., the ratio of te-form), whereas others showed improvement in both the short and long term (the ratio of complex sentences and ratio of clause conjunctions). The three sentence-ending variables showed no improvement in both the short and long term (see Tables 9.1 and 9.2). Deterioration was evident over the long term for one measure (the ratio of expressive endings).

These mixed findings may be attributable to the FL instructional context. According to Ortega (2003), students may have difficulty developing syntactic complexity in FL contexts. Based on her analyses, Ortega suggested that FL learners’ syntactic complexity might develop more slowly and less fully than learners in a second language context. FL researchers (e.g., Reichelt, 2001) have noted that a difference exists in the general learning characteristics between second language and FL settings. Second language learners have more opportunities to write
in the TL compared to FL learners, because in a second language context, students are likely to use their TL to write assignments for other subjects.

The difference in the instructional context between Japanese as a FL (JFL) and Japanese as a second language (JSL) also explains why the findings of this study differ to those reported by others investigating development in syntactic complexity. For example, Tanaka’s (1997) study, which reported that learners’ syntactic complexity (both sentence complexity and sentence-ending variables) showed improvement, was conducted with students who spent a year in Japan as exchange students (i.e., a JSL context). It is also important to note that Tanaka only reported descriptive statistics (average improvement) rather than test results for statistical significance.

It is interesting to note that only two measures (ratio of complex sentences and ratio of clause conjunction) showed improvement in both the short and long term. Improvement in the use of some measures and not others could be attributable to direct and explicit classroom instruction. For example, the use of clause conjunctions showed short-term improvement as well as a further improvement in the long term. According to the course outlines (see Appendices I and J), several grammar structures related to clause conjunction were taught in Japanese 3 and in Japanese 4. The clause conjunctions that tended to be used by the students, for example, conjunctive particles such as ga (but) and node (because), conjunctational forms such as reba-form (if), the tari-form (and) were explicitly taught in class (see Chapter 6, Section 6.3.4). The small number of conjunctions taught perhaps made these manageable for the students to remember. As Damien notes, he was able to remember the grammar patterns because he only had to learn two or three grammar patterns a week:
The grammar patterns are actually sticking in my mind…. We only learn like two or three grammar patterns a week. This is kind of study that I like.

Because there is only a few things I can concentrate on. (Damien, Japanese 3)

Although there was no instruction to use complex sentences in the individual assignments in both Japanese 3 and Japanese 4, the written work from students revealed that most of them tended to use complex sentences in their assignments. The students saw these assignments as opportunities to display what they had learned in the grammar sections of the classes that dealt with sentence patterns. For example, in Japanese 3 Jimmy commented that it was necessary to revise all the grammar that he learned during the semester to write up his individual assignment:

*The final assignment was good because that was covering the whole semester so you had to go back and learn all the grammar points and find a way to include that to the assignment.* (Jimmy, Japanese 3)

Alicia (Japanese 3) also said, ‘the individual assignment was very good for that [writing improvement]… You repeat all the grammar and it is good’. Hayashi-sensei designed the individual assignments in both subjects specifically to develop students’ writing skills. Nevertheless, both teachers in Japanese 3 commented that they thought students faced difficulties with the assignment (to create a webpage). The resources the students used may have contained different vocabularies and grammar they had not yet learned in Japanese 3. In addition, students may not have had an appropriate sample to guide them on how to write their assignment.

It is also possible that, following the explicit instruction, students noticed various complex sentences in the readings given prior to the lectures, as well as the readings that students did before preparing their assignments. Both teachers believe that appropriate input through reading allows learners to develop their writing skills.
Hayashi-sensei (the coordinator) hoped that through readings, students could learn how Japanese sentences should be structured.

Some examples from the readings students were required to do before each lecture (non-assessed) are shown below (translations are provided underneath):

Example 1:
会社が始まる9時までに東京に着くためには、朝6時に大阪を出発したら間に合わない。
If you leave Osaka at 6 am in the morning, it will be too late if you need to arrive in Tokyo before a 9 am work-commencement time.

Example 2:
今でも日本人は長い間使った道具や人形をすてないで、神社や寺へ持って行くと、おはらいをしてからもやしてもらう。
Even now, Japanese people do not throw away their used tools and dolls, they take them to shrines and temples and they are purified after being burnt.

The instruction perhaps primed student to notice complex sentences in the readings. Thus, through noticing (Schmidt, 1990) and written output practice (DeKeyser, 2007) learners were able to use a greater number of complex sentences over time. ERs can assist not only in vocabulary acquisition, but in exposing learners to different sentence structures (e.g., Hafiz & Tudor, 1990; Lai, 1993; Lee & Hsu, 2009; Mermelstein, 2015). It was evident that students noticed these complex structures. For example, Anna was aware that the readings assisted with her writing, especially with ordering sentences and with particles (one of the grammar elements used for clause conjunctions in Japanese), as she mentioned both in Japanese 3 and Japanese 4 interviews.
By reading the websites I think it helped with ordering the sentences and also with the particles (Anna, Japanese 3).

Sometimes you will see a grammar pattern in a sentence that I would have thought I would be using and things like that. And also how setting everything out, like particles and things like that like make the entire sentence (Anna, Japanese 4).

The findings also show that the te-form only improved in the short term (see Table 9.1) and no improvement was seen in the long term for this measure (see Table 9.2). According to the curriculum documents, more grammar patterns were taught that contained te-form patterns in Japanese 3 compared to Japanese 4 (8 in Japanese 3 and 3 in Japanese 4). As such, this may explain why there was an improvement in the te-form in the short term after students completed Japanese 3, whereas the post-test conducted after students had completed Japanese 4 showed no further improvement on this measure. Students may have forgotten some of the grammar related to te-form learned in Japanese 3 because they did not use them in Japanese 4.

9.1.3 Accuracy

The three measures used in this study to examine accuracy were the ratio of error-free sentences, the ratio of EFCs and words per error-free sentences. The short-term results revealed an improvement in the ratio of EFCs, with a large effect size and words per error-free sentences, with a medium effect size. However, there was no improvement in the long term and deterioration was evident in the ratio of error-free sentences.

One possible explanation for the difference in these results between the short and long term may be that the individual assignment in Japanese 3 was similar to the pre- and post-test topic. The individual assignment in Japanese 3 asked students to
produce a tourist webpage in Japanese and the pre- and post-test topic was to introduce a place the writer knew well. In contrast, the assignment topic for Japanese 4 was a research task on any writing topic related to Japanese writing. As students completed the post-test not long after they had written their assignments in Japanese 3, it may have been easier for them to write on this topic, as it was similar and fewer errors may have been made. The few studies (e.g., He & Shi, 2012; Winfield & Barnes-Felfeli, 1982) that have explored the issue of topic knowledge and writing performance suggest that students tend to perform better when given a general topic with which they are familiar. For example, in He and Shi’s (2012) study, students made more language errors when given a knowledge-specific task as opposed to a general task. Hence, improvement in accuracy may be due to topic familiarity rather than a reflection of students’ greater control over accuracy when writing on familiar topics.

One factor that may explain why few robust improvements were apparent in accuracy, particularly in the long term, was the lack of feedback. Students commented in the interviews at the end of Japanese 3 and Japanese 4 that, due to a lack of feedback from teachers, they had difficulty in improving their writing (especially in terms of accuracy) and some commented that they lost motivation. For example, Mary commented:

*If he gave me feedback, if he circled this and wrote what is wrong, or if he said you can replace with this word... that would be really useful but because he didn’t do that, it is a bit de-motivating.* (Mary, Japanese 3)

The only opportunity for students to receive feedback was when they undertook the weekly writing tasks (3 sentences), but the feedback they received was
Many students mentioned that during class, teachers only checked the weekly writing tasks briefly or did not give feedback, because of time constraints.

Some students noted they were unsure about whether they were using grammar and vocabulary appropriately. Mary (Japanese 4) admitted that ‘each week we learnt different vocab and grammar so when I use it to create a sentence I am not too sure whether it is right or wrong’. Feedback could have provided the students with that reassurance.

Previous studies (e.g., Bae & Lee, 2012; Storch & Tapper, 2009) have shown the importance of feedback for improved grammatical accuracy. Many L2 writing scholars (e.g., Bitchener & Storch, 2016; Ferris, 2003; Leki, 1990) argue that feedback is a critical element for the development of writing accuracy. When learners consciously attend to written corrective feedback input, it is likely they will notice any mismatch between their output and their TL input (Bitchener & Storch, 2016). The results of recent studies (e.g., Bitchener & Knoch; 2010, Shintani & Ellis, 2013; Stefanou, 2014) support this argument—learners who received corrective feedback were able to produce accurate revisions, as well as maintain their accuracy (in structures for which they received feedback) on subsequent occasions.

The importance of feedback for improving writing was shared by the two teachers interviewed in this study—they believed it was important for students to know whether they were writing grammatically correct sentences or not. However, the teachers also said that due to the limited time available, it was very difficult for them to provide detailed feedback to each student. Thus, teachers may believe in the benefits of feedback, but classroom realities may preclude them from providing such feedback.
9.1.4 Fluency

A number of measures were used to investigate fluency in this study. These were the number of words, KCs, KWs, sentences and clauses. Findings for both the short and long term showed that students’ writing on all measures of fluency improved (see Tables 9.1 and 9.2). The improvement was statistically significant (p < .001). There was not a great difference in terms of improvement between the short and the long term. In other words, students did not improve much more in the long term. In fact, the mean difference between the pre- and post-test was slightly greater for the short-term compared to the long-term results (but this was a small difference).

One of the ways in which fluency improves is when students develop a larger vocabulary (Nation, 2015). In both subjects, approximately half of the students commented on their growing vocabulary (see Section 9.1.1). As with improved kanji complexity, ERs during both subjects attributed to improvement in students’ vocabulary (including kanji) and thus improved fluency in both the short and long term. In both subjects, several students commented that they had more vocabulary after the semester, and having a wider vocabulary helped them write more text. For example, Norma (Japanese 4) felt that she was able to write a longer text in her post-test compared to the pre-test: ‘I think I have more vocabulary and I can write longer’. Moreover, knowing more kanji enabled students like Alice (Japanese 4) to save time and write more as they did not have to look up kanji: ‘I think it was easier for me to write because I didn’t have to go and look up the kanji, especially when handwriting’.

A number of scholars have argued for the importance of practice in promoting language development. DeKeyser (2007), in his proposed skills acquisition theory, argues that prolonged systematic practice can lead to automatised
knowledge. In other words, extensive practice decreases the learner’s reaction time and error rate. Thus, meaning-based practice can help develop fluency due to the decline in a learner’s reaction time. In both Japanese 3 and Japanese 4, students had a number of writing assignments and this too may have helped increase their vocabulary knowledge and therefore their fluency.

The findings on fluency in the current study are similar to Tanaka’s (1997), which also showed that fluency (number of words) was the area in which students made the most improvement after one year of immersion experience in Japan. In comparison with other measures (e.g., syntactic complexity, accuracy) it seems that fluency is most likely to improve. Knoch et al.’s (2015) study reported that fluency increased in ESL students’ writing after three years of study at an English-medium university, despite the lack of extensive writing required in the students’ course work. Their study suggests that fluency can improve if students are exposed to academic texts and are immersed in the language daily. In the current study, there were a number of activities and assessment tasks given throughout the semester that required ER (e.g., individual and group assignments, weekly activities from textbooks) and these readings may have assisted in improving their fluency.

A possible reason for why students did not improve further in fluency in the long term may be because students were able to write an adequate amount in the Japanese 3 post-test and it would have been difficult for them to write more in the same amount of time (i.e., 30 minutes) for their Japanese 4 post-test.

9.1.5 Content

Content was assessed on two measures: sufficiency and persuasiveness (see Chapter 6, Section 6.4.1). The main findings for each measure are discussed, along with possible explanations as to why they improved.
9.1.5.1 Sufficiency

The findings showed that sufficiency improved in both the short and long term (see Tables 9.1 and 9.2). In Chapter 6, sufficiency was determined (using a rating scale) by the amount of words and details provided by the writer. Table 7.9 outlined a statistically significant difference for changes in the short term ($M = 1.22$ for pre-test and $M = 2.20$ for post-test), as well as in the long term ($M = 1.23$ for pre-test and $M = 2.18$ for post-test). Although gains were made in both the short and the long term when compared to scores on the pre-test, there were no changes in sufficiency in the long term compared to the short term (i.e., no further improvement).

Extract 9.1 and 9.2 show the pre- and post-test scripts written by Jasmine in Japanese 3, which is an example of an improvement in sufficiency in the short term. In her pre-test, Jasmine received a score of 0 out of 3 as she only wrote a few sentences and there were not enough details about Malaysia (see Extract 9.1). In the Japanese 3 post-test, she scored 2 out of 3. As seen in Extract 9.2, Jasmine provided more information about her destination (e.g., information about the capital and the tourist attractions). Further, she gave recommendation to use public transport rather than driving, due to the bad traffic in Malaysia. Although she provided more information in her Japanese 3 post-test, each piece of information provided was brief (not elaborated); therefore, a score of 2 out of 3 was given. In her Japanese 4 post-test, she also scored 2 out of 3. Similar to her Japanese 3 post-test, although Jasmine provided sufficient information, each piece was brief. This suggests that there was no further improvement from the Japanese 3 post-test.
Extract 9.1

One Section of Jasmine’s Japanese 3 Pre-test Writing Script

Note: The examples were taken directly from the learners’ writing script. Therefore, some sentences are not necessarily written correctly. The English translation was translated as closely as possible but some errors were unable to be translated.

Score (0): Only a few sentences/words

Malaysia is a very interesting country. Since the weather is always hot, there were many customers. Malaysia’s culture is very special.
Chinese as well as Indians and Malay. It is full of shopping and delicious foods. It is cheap.

Extract 9.2

One Section of Jasmine’s Japanese 3 Post-test Writing Script

Score (2): A good range of information is included but each piece may be brief

Yes, let’s talk about Kuala Lumpur. Kuala Lumpur is the capital of Malaysia. People of Malaysia call the capital as ‘KL’.

Since KL is the capital, there are many people. During my grandfather’s era, KL was not crowded. However, the city has become very lively now due to many tall buildings as well as working people and tourists.
There is a world famous KL twin tower in KL. That is Malaysia’s tallest building. If you are a tourist that likes to go shopping, KL is a good place because there are many shopping malls and it is cheap to shop.

Driving in KL is not convenient because there is too much traffic. It is better to use either a bus or a train.

If you come to KL, you don’t need to worry because most of the things are cheap. You only need RM 100 ($40) for one day.

Improve in sufficiency can be attributed to the individual assignment, especially in Japanese 3. The genre of the pre- and post-tests was the tourist destination review (introducing a place). In Japanese 3, students were exposed to this kind of genre throughout the subject. It is likely that students encountered the models/words used in their post-tests through the various activities during the semester, and thus were able to use them well. This then led to improvements in sufficiency. For example, as part of their individual assignment in Japanese 3, all the participants (N = 41) wrote a description of how to reach a chosen Japanese tourist location (e.g., from Tokyo to the chosen destination). Some examples from participants’ assignments are shown below (Extracts 9.3 and 9.4). Mary (Extract 9.3) described how to reach Hiroshima from Tokyo station, including what number ‘Shinkansen’ (high-speed railway) to take. Ricki provided a very detailed description (see Extract 9.4) on how to reach Kinkakuji temple, starting from Narita airport to Kyoto station, eventually arriving at Kinkakuji temple.

Extract 9.3

One Section of Mary’s Individual Assignment

どうやって広島にいきますか。
東京駅から、105番の新幹線‘のぞみ’に乗って、広島駅まで行きます。料金は 18880 円です。4時間ぐらいかかります。

How do you get to Hiroshima?
From Tokyo station, take the 105 high-speed railway 'Nozomi' to Hiroshima station. The cost is 18,880 yen. It takes approximately 4 hours.
Extract 9.4

*One Section of Ricki’s Individual Assignment*

Transportation guide to Kinkakuji Temple

*Firstly, take the Narita express from Narita airport to Tokyo station. Then, take the high-speed railway from Tokyo station, 2 hours later you will arrive at Kyoto station. Then, it is about approximately one minute walk to the bus station. Take the 205 bus and get off at Kinkaku michi bus stop. You will arrive at Kinkakuji Temple in approximately 40 minutes.*

It was interesting to observe the transfer from this assessment task to the post-test. Many students (28 out of 41) included information about the time taken to reach a certain destination and the modes of transport in their post-test. Extracts 9.5 and 9.6 are representative of the students’ writing, showing the inclusion of information about the mode of transportation to reach another location. For example, David included the mode of transportation (i.e., tram) to explain how to reach Frankston station from Flinders station (see Extract 9.5). Similarly, Sharon (see Extract 9.6) explained two modes of transportation (i.e., plane and large car) to reach Yunnan and further described the way of travelling to Yunnan from Melbourne (e.g., via a transfer in Hong Kong). Thus, the genre practiced in Japanese 3 may have assisted students in their post-test.

Extract 9.5

*One Section of David’s Japanese 3 Post-test Writing Script*

*St Kilda is situated in the south of Melbourne. It takes about 15 minutes by tram from Flinders Station to Frankston station. You will get there by taking the No.98 tram.*
In Japanese 4, the genre taught and practiced was formal written correspondence. It was not related to the pre- and post-test genre (i.e., tourism). This may explain why there was no further improvement in sufficiency following Japanese 4.

9.1.5.2 Persuasiveness

The findings revealed that persuasiveness improved in the short term as well as in the long term (see Tables 9.1 and 9.2). As explained in the methodology section (Chapter 6), persuasiveness was examined by determining whether the writer included reasons for why somewhere was worth visiting and/or whether that place featured something unique or special. Similar to sufficiency, there was a statistically significant difference for changes in the short term (M = 1.17 for the pre-test and M = 1.98 for the post-test), as well as in the long term (M = 1.14 for the pre-test and M = 1.91 for the post-test). Although there were gains in both the short and long term, there were no changes in persuasiveness in the long term compared to the short term (i.e., no further improvement from the short to the long term).

Extracts 9.7 and 9.8 show the Japanese 3 pre- and post-test scripts written by Helena as an example of improvement in persuasiveness in the short term. In her pre-test (see Extract 9.7), Helena explained some activities that could be done in
Melbourne (e.g., being able to eat different kinds of food), but the reasons for visiting Melbourne were limited; therefore, a score of 1 out of 3 was given. In her post-test (see Extract 9.8) however, she scored 2 out of 3. In her post-test, Helena begins by asking the audience whether they would like to come to Mt Fuji. This was different from her pre-test, as she only made a statement, ‘It is Melbourne’, whereas in her Japanese 3 post-test she used a questioning format, which includes subtle persuasion. Helena then provided five reasons for visiting going Mt Fuji. For example, she wrote that five beautiful lakes are visible from the top of Mt Fuji. Nevertheless, each reason was very brief and hence a score of 2 out of 3 was given. Helena’s Japanese 4 post-test was similar to that of her Japanese 3 post-test. The same score was given, which meant that there was no further improvement—Helena provided reasons for visiting Melbourne in her Japanese 4 post-test; however, they were also brief. Even some of the expressions she used in the Japanese 3 post-test and Japanese 4 post-test were similar. For example, the expression ‘you can do a lot of enjoyable things’ that appeared in her Japanese 3 post-test also appeared in her Japanese 4 post-test.
One Section of Helena’s Japanese 3 Pre-test Writing Script

Note: The examples were taken directly from the learners’ writing script. Therefore, some sentences are not necessarily written correctly. The English translation was translated as closely as possible but some errors were unable to be translated.

Score (1) Gives appropriate reasons for visiting the place but may be limited to one or two.

It is Melbourne. You can listen to music. You can eat various foods. You can go there by train. I like to go to Hanaback. The city is close to the beach and beautiful. Let’s go to Yarra Mountain. You can buy shoes. You can get there by plane. Let’s go to Melbourne.

Score (2) Gives appropriate and an adequate number of reasons but the information may be brief or is not elaborated further.

Would you like to come to Mt. Fuji? You can catch a bus from Tokyo to Mt. Fuji. It is about 2 hours and a half. When you go there, you can climb the mountain. Mt. Fuji is very beautiful. It is Japan’s symbol of beauty.

You can do many fun things near Mt. Fuji. For example, there is a restaurant that sells delicious food. Moreover, they sell souvenirs for your family.

At the top of Mt. Fuji you can see five very beautiful lakes. You can walk the trail. It is a very good exercise isn’t it?

Everybody, let’s travel to Mt. Fuji!
As mentioned previously, the pre- and post-tests were in the genre of tourist destination reviews, and these are promotional texts. Promotional texts are persuasive (Thumvichit, 2016), as the writer must persuade readers to visit the attraction. Several assessment tasks in Japanese 3 may explain the improvements in persuasiveness over the short term. These tasks included individual assignments (making a webpage of a tourist destination in Japan), an oral assessment task and readings. The individual assignment in Japanese 3 was to make a webpage of a tourist destination in Japan. That this task was to promote the place students had chosen (from any tourist destination in Japan) provided them with opportunities to practice writing a persuasive piece. As Ryan (Japanese 3) mentioned, when he was writing the individual assignment, he had an audience in mind that he wanted to communicate to, persuading them to visit: ‘I was writing towards the audience. And I have to bring them in and I have to communicate with them so it is sort of like talking to others’. In order to produce this webpage, students used Japanese tourist websites (of the chosen destination) as their resources. This means that students were exposed to promotional genres and examples of persuasive language. Perhaps students were able to ‘notice’ some of the persuasive language used in the resources and were then able to apply them.

The oral assessment in Japanese 3 was also a promotional genre. In this task, students were required to make a 5-minute video in a group introducing a place in Melbourne and persuading the audience to visit. This task again provided them with practice in producing persuasive expressions.

The reason for no further improvement being evident over the long term is perhaps due to the nature of tasks in Japanese 4. As these were not related to the
promotional genre, students did not have the opportunity to practice writing a persuasive piece or practice speaking persuasively.

9.1.6 Structure

Another feature of L2 written composition addressed in this study through rating scales was structure (see Chapter 6). Structure was divided into coherence and paragraphing. The findings revealed a statistically significant improvement in structure in the short term as well as in the long term (see Tables 9.1 and 9.2). The mean difference (i.e., pre- and post-test scores) over the short term for coherence was 0.76, with 1.36 for the long term. The mean difference over the short term for paragraphing was 0.68, with 1.04 for the long term. This suggests that structure improved further over the long term.

Extracts 9.9, 9.10 and 9.11 show the short-term and long-term improvements by Peter as an example of improvement in coherence. As seen in Extract 9.9, in his Japanese 3 pre-test (0 out of 3 was given), there is no connection from one sentence to another and the entire text is difficult to follow. For example, after his second sentence where he suggests that various tourist attractions and museums exist, he then mentions that Melbourne is located in the south of Australia. Moreover, due to the limited writing, there is no paragraphing at all. Peter improved slightly in his Japanese 3 post-test (see Extract 9.10). His writing still showed limited organisational flow; therefore, a score of 1 out of 3 was given. For example, in his second paragraph, Peter mentions shopping, but in the following sentence, he discusses museums. In contrast to the first two pieces of writing, Peter was able to connect the information from one sentence to another, as well as from one paragraph to another in his Japanese 4 post-test (see Extract 9.11). For example, in the second paragraph, after saying that there were many activities available in Melbourne, Peter
introduced various places for those who enjoy nature, sports and shopping. This shows a further improvement from the short to the long term.

Extract 9.9

One Section of Peter’s Japanese 3 Pre-test Writing Script

Note: The examples were taken directly from the learners’ writing script. Therefore some sentences are not necessarily written correctly. The English translation was translated as closely as possible but some errors were unable to be translated.

Score (0): Entire text difficult to follow

interesting Melbourne

Melbourne is very interesting isn’t it? There are various tourist attractions and museums. Melbourne is in the south of Australia. You shop in Melbourne, watch movies, swim at the beach and get on the tram. It is expensive to travel in Melbourne but the museum is cheap.

Extract 9.10

One Section of Peter’s Japanese 3 Post-test Writing Script

Score (1): Presents information with limited organisational flow

beautiful Melbourne

Melbourne is a city in the south of Australia. There are many tourism resources in Melbourne. When you go sightseeing, you will go to the CBD but go outside the city. There are museums, the zoo and many beautiful parks in the CBD. There is Philip Island and a big
sea road outside the city.

There are many shops in the CBD and you can go shopping. If you like museums, you go to the Melbourne museum and CBD. Melbourne museum costs $16.

Extract 9.11

One Section of Peter’s Japanese 4 Post-test Writing Script

Score (3): Main point is clear, information is organised logically; ideas are smoothly linked throughout the text

I come from a city called Melbourne. Melbourne is situated in the south of Australia and the city is beautiful. You can come to my city.

There are many places you can go and many things you can see in Melbourne. If you like nature, you can go to places like the beach, zoo and Phillip Island. If you enjoy sports, you can go to the stadium to see the tennis, football and cricket. If you like shopping, there are many shopping centres and department stores where you can shop.

How do you go to Melbourne? Usually foreigners use the plane to go to the airport, from the airport to the city you catch a taxi but you can also get on the limousine bus. In the near future they will build a train.

Melbourne is not cheap, when you travel, carry a lot of money!

That coherence showed a significant improvement in both the short and long term might be attributed to the readings and individual assignments. Studies on ER have shown that it assists improvements in structure and organisation (e.g., Lee & Hsu, 2009; Shen, 2009). This suggests that the more exposure learners have to their TL, the higher the probability of improving their writing.
Explicit teaching can also explain improvement in coherence. Hayashi-sensei’s main aim for the individual assignments in both Japanese 3 and Japanese 4 was to develop students’ ability to write a coherent essay in Japanese. As mentioned in the observation section (Chapter 6, Section 6.3.4), the concept of coherence was taught explicitly in one of the classes in Japanese 3 (Week 6). Figure 9.1 shows one of the slides (fours slides in total) used in class to teach students about coherence. A short workbook exercise was also given to practice coherence. This explicit instruction may have assisted students with improved structure.

Prior to writing their assignments in Japanese 3 and Japanese 4, Hayashi-sensei also briefly described how to organise an essay (i.e., introduction, main body and conclusion). Hayashi-sensei believed that teaching students how to structure an essay was important. He tried to focus more on teaching how to structure a research essay: ‘I think I have put more emphasis on writing a structured essay, talking about the structures of writing’. Schmidt (2001) suggests that awareness plays an important role in language learning. It is likely that students became more aware of organisation (structure) after receiving such instruction, as evident in Matt’s
(Japanese 4) comments in the end-of-year interview. He became aware of the importance of structure and planned his essay:

Yes. Probably incorporating a more holistic approach to the task instead of thinking which sentence to write. I write this sentence, what else can I write? More planning ahead and planning the whole essay.

Hayashi-sensei’s aim was for students to write coherently at the end of Japanese 4:

The goal for both Japanese 3 and Japanese 4 combined was to be able to write an academic essay. Web design was just a practice... webpage can be just a compilation of different pieces of information whereas an academic essay has to be a coherent piece of work so yes, Japanese 3 was designed to be a practice.

Research on writing development (e.g., Bae & Lee, 2012; Storch & Tapper, 2009) suggests that focusing instruction on a particular area (e.g., structure, vocabulary) will likely lead to improvement in that area. For example, Bae and Lee’s (2012) study found a significant improvement in their ESL students’ coherence following instruction and practice. In their study, the general concepts of coherence, including introduction, main body and conclusion were taught in class, and were then applied through different activities. The same was evident in this study—the general concepts of coherence were taught. These findings show that instruction plays a vital role in improving structure, as students practiced writing in Japanese 3 and 4.

Frequent readings would also have helped students with structure more generally, including paragraphing. The readings that students were required to do before each lecture in both Japanese 3 and Japanese 4 (non-assessed) contained four
to five paragraphs. Other readings that they have done throughout the semester (e.g., readings prior to writing their assignment) may also have assisted in their organisation. This is illustrated by a comment from Anna (Japanese 3): ‘by reading the websites I think it helped with ordering the sentences’.

Even though limited teacher feedback was available (in both Japanese 3 and Japanese 4), feedback from peers may have contributed to the improvement in students’ structure, especially in the long term. In Japanese 4, students engaged in peer feedback activities on their individual assignments. Hayashi-sensei included peer feedback activities in the classes, as he believed that peer feedback worked better in improving content and structure than grammar: ‘I think students were able to correct each other’s structure and contents rather than grammar’. Students’ comments support this teacher’s belief. Almost all (19) students in Japanese 4 indicated that giving feedback on the individual assignment helped with their writing, especially with structure, as they were able to reflect on their own writing. For example, Cindy (Japanese 4) said, ‘when I was commenting on other people’s structure, I would think about my own essay, what did I do and what I didn’t do’. The findings concerning peer feedback and improvements in structuring concur with those reported in studies on the benefits of peer feedback. For example, Yang et al. (2006) and Ruegg (2015) found peer feedback beneficial for improving structure.

9.2 Research Question 2

Is there any difference in L2 writing development between character-based learners and non-character-based learners?

This section draws upon the findings in relation to the second RQ. It offers a discussion about whether a difference existed in L2 Japanese writing development between character-based learners and non-character-based learners, based on the
quantitative findings. Participants’ perspectives on writing development are also discussed. Only the participants from Japanese 3 (N = 41) were used to answer the second RQ, as the number of participants in Japanese 4 was too small (N = 22).

The findings reveal that the measures with statistically significant differences in development for the two groups were mainly in kanji-related areas: kanji complexity (i.e., KC RTTR and CTTR, KW RTTR and CTTR) and fluency (number of KCs and KWs). Character-based learners improved more than the non-character-based learners between the pre- and post-tests. For example, character-based learners had a pre-test/post-test mean difference of 1.69 for KC RTTR, whereas the non-character-based learners had only 0.86. Similarly, the character-based learners improved their mean number of KWs by 26.44, whereas the non-character-based learners only improved by 11.48.

That the character-based learners showed greater improvement over one semester than the non-character-based learners may be due to these learners having familiarity with more than one character-based language. Background questionnaires revealed that half of the character-based learners (8 out of 16) were learning Japanese as their third language and the remaining half were learning it as their fourth. For example, learners who were not from mainland China (e.g., Hong Kong, Malaysia, Taiwan), were able to speak languages other than Mandarin (e.g., Cantonese, Malay, Taiwanese). In contrast, more than half (15 out of 25) of the non-character-based learners only spoke English. According to Ortega (2009), knowledge of more than two languages accelerates learning in additional ones (Ortega, 2009). The benefits are even greater if the L3 is related to other languages that the learner already knows (Ortega, 2009). Here, this explains why it was easier for character-based learners to learn kanji compared to the non-character-based learners.
That there was a difference seen only in the *kanji* between the two cohorts was because Chinese (Mandarin) and Japanese do not have a similar grammatical structure. Therefore, in terms of sentence ordering for example, character-based learners did not have an advantage over non-character-based learners.

The above findings (i.e., that background learners have an advantage) corroborates some of the past literature (e.g., Aoki, 2010; Machida, 2001; Scarino et. al, 2011) but not others (e.g., Iwashita, 2012; Ke, 1998). In Aoki and Machida’s studies, BS learners performed better in reading and writing *kanji* compared to the non-background (NBS) speaker learners. In Ke’s (1998) study (investigating L2 learners of Chinese) however, language background did not have a significant influence on a learner’s ability to recognise and produce Chinese characters. Ke suggested that the non-background learners in his study might have been highly motivated learners who invested more time learning Chinese characters compared to the background learners. However, as the information on participants’ academic profiles and their motivation for TL learning was not available, Ke could not provide an exact reason to explain these differences.

Another study that did not corroborate the findings of the current study was Iwashita’s (2012) study. Unlike the current study and other studies (e.g., Aoki, 2010; Machida, 2001) where there was a difference (mainly in *kanji*-related areas) between the two groups, Iwashita reported that the BS learners outperformed the NBS learners in all areas (content, vocabulary, form and structure, discourse, use of scripts and combination). Moreover, the feature that clearly distinguished the two groups was the rich content that BS learners produced using a variety of forms and structure. The researcher attributed her findings to the fact that BS learners had more familiarity with Japanese culture and were able to include more cultural information.
The attitude survey in Iwashita’s study revealed that BS learners frequently visited Japanese websites and engaged with Japanese material through comic books. This allowed them to become familiar with Japanese culture and they were able to include Japanese culture-related information in their writing task.

This was not the case in the current study, where there was no significant difference in development between the two groups in terms of content. The difference in the findings between the current study and Iwashita’s study may be a result of the difference in participants, as well as their learning environment and the tasks they had to complete. The participants in the current study were adults (university students), whereas participants in Iwashita’s study were secondary school students (Year 10); there may be a difference in the cognitive development between adults and adolescents. Moreover, in the current study, participants were all from the same course (i.e., Japanese 3) and thus both groups were taught the same content. In contrast, the participants in Iwashita’s study came from various schools across Australia. They were not taught using the same curriculum. The writing task in the current study was related to the curriculum; in Iwashita’s study, it was not necessarily related to the curriculum. Students came from different schools and this may have affected the results. In Iwashita’s study, Task 1 was to write a self-introduction as part of a scholarship application for an exchange program, and Task 2 was to respond to a blog sent by a Japanese student.

The quantitative results showed that character-based learners improved more compared to non-character-based learners in kanji and teachers were aware that character-based learners had an advantage when learning kanji. However, they believed that ultimately improvement depended on each individual rather than the
background. As Hayashi-sensei noted, ‘in terms of learning kanji and writing kanji, yes it has to do with each individual and their own experiences of learning Japanese’.

Despite it being time consuming for non-character-based learners to learn kanji, it seems they still enjoyed it. Those students who said they enjoyed learning kanji were all non-character-based learners (character-based learners neither enjoyed nor disliked learning kanji). For example, Christiana (Japanese 3), a non-character-based learner enjoyed learning kanji because they are like pictures: ‘I find it interesting, trying to make picture out of it. I tried to make a picture with all of them, yes I enjoy it’. Moreover, most of the non-character-based learners did not find it a struggle to learn kanji. Students like Anthony even thought that kanji learning was easier than learning other areas such as grammar:

_Yeah I think relatively easy compared to like memorising the grammar points._

_Because with the grammar points I need to actually write the sentences and do the assignments to learn them. But with the kanji it just goes into my head._

_(Anthony, Japanese 3)_

According to Kormos (2012), writing is a time-consuming and complex task that requires determination and concentration. As such, a learner’s motivation and self-regulation are important in determining whether he or she will ‘engage in writing activities, what kind of writing task they will undertake, with what level of effort and attention they approach the various phases of the writing process, and how they exploit the learning potential of [the] writing task’ (Kormos, 2012, p. 2). This can also be said when learning kanji, especially for non-character-based learners—if they are motivated and find it interesting, they will devote the time to kanji learning. For example, Corinna (Japanese 3) spent five hours a week learning the required 10 kanji: ‘Five focused hours throughout the week, yes. Or it may take more. I guess it
depends. Each *kanji* 25 times’. Even though it took her a long time and she found them difficult to remember, she still found it interesting: ‘I enjoy it. I find memorising the *kanji* itself is the tricky part because you have to associate it with the reading and remember it day to day you have to be using it’. Thus, positive self-efficacy and interest are particularly important for *kanji* learning.

Student interviews also revealed a difference in *kanji* learning strategies between the character-based learners and non-character-based learners. The majority of character-based learners tended to focus on the difference in *kanji* and Chinese characters. For example, Anne (Japanese 3) commented that it was sometimes difficult to tell the difference between the two: ‘the difference between Chinese and Japanese characters. Like one stroke difference, it is hard to pick it up but when the teacher says I realise’. Similarly, Jasmine (Japanese 3) said, ‘sometimes when the *kanji* is similar to Chinese, I have to make sure not to get confused. I will have to write out a few times for those ones’. Further, in the teacher interview, Tanaka-sensei commented that character-based learners tended to put less effort into *kanji* learning because of their background and they end up making errors.

In contrast, non-character-based learners needed to memorise *kanji* rather than remember the difference between *kanji* and Chinese characters. With the increase in the use of technology, some non-character-based learners have commented on the usefulness of applications and computer software for *kanji* learning. For example, as Peter (Japanese 3) mentioned, ‘I use a software called anki—I create a deck. Use this to learn *kanji*’. Ben also noted that he used an application to learn the *kanji* for the weekly test:

*For me I have been using an application to learn kanji I am better at recognising kanji more than writing, it is called wanikani. Usually I practice*
the day before [i.e. before the kanji test]. With the programs that they use, recognising them is not too difficult and after that writing is not too difficult.

(Ben, Japanese 3)

The findings reveal that character-based learners showed greater development compared to the non-character-based learners, but only in terms of learning kanji. Other than the kanji-related measures, no statistically significant differences existed between the two groups on the wide range of measures used in this study to assess development. This lack of difference between the two groups could be because both groups received the same instruction.

9.3 Study 2 Summary

The discussions in the previous sections provided possible explanations for the findings in Study 2 in relation to RQs 1 and 2. Overall development in writing was evident, as most of the measures investigated showed improvement. In both the short and long term, most measures showed improvement (other than a few measures in accuracy and syntactic complexity). The findings also reveal that the difference in development between the character-based learners and non-character-based learners were only in kanji-related measures.

Various activities (assignments, readings, weekly kanji tests and answering textbook questions) and instruction (weekly grammar instruction and instruction on structuring an essay) resulted in students’ writing development. Syntactic complexity and accuracy may be more difficult to improve in compared to other measures, especially with insufficient feedback. The findings suggest that instruction is critical for writing improvement.

Following is a discussion of the Study 1 and Study 2 results of the current project.
9.4 Discussion of the Entire Project (Study 1 and Study 2)

In the current project, writing development was investigated using a cross-sectional study (Study 1) as well as a longitudinal study (Study 2). The writing features used in Study 1 (and that were applied in Study 2) were kanji complexity, syntactic complexity, accuracy and fluency. In this section, the outcomes of Study 1 and Study 2 are compared and discussed. Table 9.3 summarises the developmental changes seen in students’ writing.

Table 9.3

Summary of Developmental Changes in Study 1 and Study 2

<table>
<thead>
<tr>
<th>Writing features</th>
<th>Study 1 (cross-sectional)</th>
<th>Study 2 (longitudinal)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kanji complexity</td>
<td>As proficiency increased, <em>kanji</em> complexity (e.g., KC RTTR) also increased. Only the upper level did not show significant differences</td>
<td>Improved both in the short and long term</td>
</tr>
<tr>
<td>Syntactic complexity</td>
<td>Mixed results. <em>Ratio of clause conjunction:</em> only measure that showed significant differences between all levels. <em>Ratio of complex sentences:</em> showed slow increase at wider intervals. Other measures in general only showed significant differences in the lower proficiency level.</td>
<td>Measures that showed improvement both in the short term and long term: <em>Ratio of complex sentences</em> <em>Ratio of clause conjunction</em> Te-form (short-term improvement only)</td>
</tr>
<tr>
<td>Accuracy</td>
<td>Ratio of EFCs: steady development</td>
<td>Improved only in the short term <em>Ratio of EFCs</em> <em>Words per error-free sentence</em></td>
</tr>
</tbody>
</table>
As seen in Table 9.3, similarities and differences are both evident in the developmental changes in Study 1 and Study 2. These are discussed below.

The writing features that showed similar developmental patterns in both studies were kanji complexity and fluency. In Study 1, there was a positive linear relationship between the five proficiency levels for kanji complexity and fluency (except between Levels 7 and 9 for kanji complexity). In Study 2, these two writing features improved in both the short and long term. The findings from Study 1 and Study 2 suggest it is likely that learners can improve their kanji complexity and fluency as their proficiency increases, but that once students reach a high level (i.e., Level 7 in this study), it may be more difficult to improve. This may explain why the participants in Study 2, who were only in the lower-intermediate level (Levels 3 and 4) showed improvement in kanji complexity (KC and KW) in both the short and long term. Fluency, on the other hand, appears to show the clearest signs of improvement in writing compared to other features. This has been reported in a number of studies investigating writing, and this improvement occurs regardless of whether the L2 is learned as a foreign (e.g., Tanaka, 1997) or second language (e.g., Knoch et al., 2015), as learners’ vocabulary knowledge increases.

The results for syntactic complexity showed mixed results. For example, the ratio of clause conjunctions showed significant differences between all levels in Study 1 and this measure also improved in both the short and long term in Study 2. The ratio of complex sentences showed a slow increase at wider intervals in Study 1 (i.e., significant differences in the lower levels then levelling off and again showing an increase) and showed improvement in Study 2 in both the short and long term.
The three sentence-ending variables (propositional endings, epistemic endings and expressive endings) showed significant differences between Level 1 and other levels in Study 1 (i.e., lower proficiency level) but did not show any changes in Study 2 in the short or long term (lower-intermediate levels).

One possible factor that may explain these varying results in syntactic complexity both in Study 1 and Study 2 (i.e., that some measures improved more than other measures) may be that students were explicitly taught how to use some writing structures (e.g., clause conjunction) while others were not explicitly taught (sentence-ending variables). As seen in Study 2, curriculum documents in Japanese 3 and Japanese 4 included clause conjunctions as part of the grammar instruction. Following these explicit instructions, students were able to ‘notice’ the syntactic complexity taught and could then apply and ‘practice’ them when writing their assignments. Another possible factor that may explain these results may be the writing tasks that students were asked to do. The task in Study 1 was to write an email to a Japanese-speaking friend in Japan about a trip to Japan. The task in the pre- and post-tests in Study 2 was to introduce a place the writer knew well. Both tasks were descriptive and there may have been limited opportunities to use sentence-ending variables to show development. Finally, as discussed earlier (see Section 9.1.2), syntactic complexity may be difficult to improve in, particularly for students learning their L2 in a FL context (Ortega, 2003). Thus, some measures may improve at a lower-intermediate level, but this may not be necessarily the case with other measures unless students have the opportunity to use and practice them frequently.

Accuracy was one writing feature that showed a difference in the results between Study 1 and Study 2. The ratio of EFCs showed a steady development in
Study 1; however, this measure only improved in the short term in Study 2. Further, the other two measures only showed significant differences in the upper proficiency levels in Study 1. Study 1 was a cross-sectional study and no qualitative data (i.e., interview results) was available to determine the reason behind development or the lack of development. However, the Study 1 results, as well as past studies (e.g., Cumming et al., 2006; Knoch, Macqueen et al., 2014) suggest that higher-proficiency level learners tend to write more accurately. In Study 2, one possible explanation for why accuracy improved only in the short term was because the individual assignment in Japanese 3 was similar to the pre- and post-test topic. Thus, it may have been easier for students to write on a topic that they had ‘practiced’ before. As suggested by DeKeyser’s (2007) skills acquisition theory, the ‘power law of practice’ leads to automatised knowledge and assists in improving accuracy. Further, one factor that may explain why no improvements were evident in accuracy—even in the long term in Study 2—was the lack of feedback (as reported in student interviews). As mentioned earlier, feedback is critical for improving learners’ accuracy (e.g., Bae & Lee, 2012; Storch & Tapper, 2009).

The findings from Study 1 and Study 2 suggest that kanji complexity and fluency are likely to improve as a learner’s proficiency increases. In contrast, syntactic complexity and accuracy are difficult to improve in compared to the other two writing features. In addition, some measures for syntactic complexity (e.g., Japanese-specific sentence-ending variables) may not show developmental changes compared to other measures, such as the ratio of clause conjunctions. Thus, further studies on Japanese-specific measures may be necessary. Past research (e.g., Knoch et al., 2015) notes the importance of qualitative data (i.e., interview and observations). It enables researchers to determine the reason behind improvements or
the lack of improvement. This project also suggests that qualitative data are critical to gain information on what factors lead to learners’ writing development or the lack thereof.

Further, the findings here reinforce the importance of a longitudinal study. Longitudinal studies are invaluable as they allow researchers to study learners over an extended period. For example, learners may not necessarily improve their writing in a positive and linear way, as evident in this study. Some of the measures in accuracy only improved in the short term. It is important to know why, and this may sometimes be revealed through interviews.

Moreover, depending on the context of the study (e.g., classroom-based) improvements in learners’ writing may differ. In addition, the current project suggests that although teaching writing is considered important for character-based languages (Cook & Bassetti, 2005), there is still little emphasis on writing, as shown in Study2. Minimal emphasis on writing in class may be one reason behind the lack of improvement in some writing features (accuracy and syntactic complexity).

In the final chapter of this thesis, I will draw on the main findings of both Study 1 and Study 2, the findings of the overall project and I will outline some directions for further research.
Chapter 10: Conclusion

The objective of this research project was to investigate the writing development of L2 learners of Japanese at an Australian university. This research project was divided into two studies: a cross-sectional study (Study 1) and a longitudinal classroom-based study (Study 2). In Study 1, the main aim was to examine how the writing ability of university learners of Japanese differed across proficiency levels. Study 2 aimed to investigate the writing development of lower-intermediate L2 Japanese learners, using some of the measures from Study 1, and supplementing them with rating scales. Qualitative data in the form of interviews with students and teachers and classroom observation notes were used to explain the findings.

This research project was motivated by a lack of research on writing development in LOTE. Previous research on L2 writing development has been conducted mostly in either an ESL or EFL context. Further, these studies are either cross-sectional (e.g., Cumming et al., 2006; Verspoor et al., 2012; Knoch et al., 2015) or longitudinal (e.g., Knoch et al., 2014; Storch & Tapper, 2009; Sasaki, 2007, 2009, 2011). The current research project combined a cross-sectional and a longitudinal design. The findings from Study 1—the cross-sectional study—informed the measures deployed in Study 2. Theoretically, Study 2 was informed by two cognitive theories of SLA—the noticing hypothesis (Schmidt, 1990) and the notion of practice in skill acquisition theory (DeKeyser, 2007). The noticing hypothesis can explain improvement in writing by reference to interventions (e.g., feedback) which can help learners ‘notice’ the gap in their knowledge. The notion of practice in skills acquisition theory suggests that through meaning-based practice,
learners are able to automatise their knowledge and thus become more fluent and accurate in their writing.

In this chapter, a summary of the key findings will be provided first, followed by the theoretical, methodological and pedagogical implications of the research project. The limitations of the study and suggestions for future research will also be addressed.

10.1 Summary of Findings

This section first presents the findings from Study 1, followed by Study 2 findings. The overall findings are then presented.

10.1.1 Study 1 findings

Study 1 set out to investigate how the writing ability of L2 learners at university differed across levels. This was particularly important, as few studies have used such a variety of measures as in this study when investigating the writing of learners other than in English. A total of 146 writing scripts from students who took part in a Japanese placement test in a large metropolitan Australian university were used as data and were analysed using a range of discourse measures. The data analysis was quantitative.

In Study 1, four measures were used for kanji complexity, six measures for syntactic complexity, three measures for accuracy and five measures for fluency. Overall, the measures of kanji complexity and fluency showed more developmental changes compared to measures of syntactic complexity and accuracy. The findings suggest that the syntactic complexity and accuracy measures used in this study were not able to capture the developmental changes for all the proficiency levels. In general, syntactic complexity showed significant developmental changes between
the lower proficiency levels and significant differences were evident in the upper proficiency levels for accuracy.

As Study 1 was a cross-sectional study, it could not observe the writing development of the same group of students over time. Thus, Study 2 used a longitudinal mixed-methods design to examine the writing development of the same group of students.

10.1.2 Study 2 findings

Study 2 investigated the writing development of lower-intermediate L2 Japanese learners over two periods: a short-term (one semester) and long-term (one year i.e., two semesters) one. A pre- and post-test design was used and interviews were conducted with both students and teachers in both semesters. A mixed-methods approach was used to analyse the data. Overall, a comparison of the pre- and post-test results showed development in both the short term and long term. This analysis showed that the students’ writing improved in complexity, fluency, accuracy, content and structure. However, this improvement was not evident in all the measures used. Several measures of syntactic complexity and accuracy showed no change in the short and long term or only showed improvement in the short term.

An analysis of the qualitative data (interview data and observation notes) suggested that several factors might have contributed to the learners’ writing development. The most important factor that can explain these findings is the ER that students had to undertake for their individual assignments and the weekly kanji tests, which encouraged memorisation. The readings and weekly kanji tests led to vocabulary acquisition and reflected the gains in measures of kanji complexity, fluency and content.
Other researchers (e.g., Lee & Hsu, 2009; Mermelstein, 2015), investigating the writing skills of EFL learners have also shown that reading leads to vocabulary acquisition. What perhaps distinguished the reading that students in this study undertook was that it was for a purpose; that is, for completing assignments. This encouraged the learners to notice new vocabulary they could use in their assignments. This was made clear in the students’ interviews. According to the noticing hypothesis (Schmidt, 1990, 2001), language forms must be noticed before they are acquired. Nation (2001) concurs by suggesting that reading encourages learners to notice new words and remember them. Despite the concerns that Hayashi-sensei (the coordinator of the Japanese3/4 cohort) had about the assigned reading being too difficult for the students, more than half of the students commented that reading Japanese resources assisted with their writing. Thus, the findings reinforce the importance of reading for learning vocabulary and subsequently for writing development.

Another factor that can explain the development in students’ writing in the short and long term was explicit instruction. Past studies on writing development (e.g., Bae & Lee, 2012; Storch & Tapper, 2009) have shown that explicit instruction in a particular area (e.g., structure, vocabulary) leads to improvement in that area. Explicit teaching, such as weekly grammar instruction in Japanese 3 and Japanese 4, made students conscious of the various grammatical patterns and structures. It may have also primed them to notice these language forms when they came across them in their reading. The students then practiced those new grammar patterns in their assignments. According to DeKeyser (2007), meaning-based practice will reduce reaction time and error rates—a quality known as ‘power law of practice’ in skill acquisition theory (p. 3).
Although overall improvement in learners’ writing was evident, some measures for accuracy and syntactic complexity did not show improvement. Lack of development on accuracy may be explained by the students receiving insufficient feedback. A number of L2 writing scholars (e.g., Bitchener & Storch, 2016; Ferris, 2003; Leki, 1990) suggest that feedback is important for the development of writing accuracy. It was evident from the class observation and acknowledged by the teachers that due to limited time, very little teacher feedback was given on students’ writing during class. The students also noted the absence of teacher feedback on their written classroom writing tasks and commented about the difficulty of improving their accuracy without such feedback. Thus, the findings suggest that not only ERs and explicit instructions are important for writing development in the long term; repeated feedback is also an important factor.

Study 2 also considered whether character-based learners showed greater improvement in their writing development than non-character-based learners, and if so on what aspects of writing. Examining the difference in development between the character-based learners and non-character-based learners is an important issue, especially in countries like Australia where the number of international students has increased exponentially in recent years. The majority of these international students have backgrounds in other Asian character-based languages and are often studying Japanese alongside those from English or other backgrounds (non-character-based learners). Background learners are assumed to have an advantage over non-background learners.

In the current study, the differences in development between the character-based learners and non-character-based learners were only noticed in kanji-related measures. Other measures did not show any statistically significant differences.
between the two groups. This was contrary to Iwashita’s (2012) study where it was reported that BS learners outperformed the NBS learners in all areas. The difference in the findings between the current study and Iwashita’s study may be due to the difference in participants’ background, their learning environment and the tasks that they had to complete.

10.1.3 Findings of the entire project

The findings from Study 1 and Study 2 showed that learners are likely to improve their kanji complexity and fluency as their proficiency increases; however, with kanji complexity it may be difficult to show improvement when they reach a high level. Compared to other features of writing, fluency seems to exhibit the clearest signs of improvement in writing, as was also seen in other studies (e.g., Knoch et al., 2015). However, the findings for syntactic complexity showed mixed results in both Study 1 and Study 2: some measures improved while others did not. The Japanese-specific measures (i.e., Japanese-specific grammatical structures) showed less developmental changes compared to other measures. Perhaps the mixed results may be because some of the grammatical structures were explicitly taught in class and students were able to practice them and improve, which then was reflected in some of the measures investigated. Another possible reason may be due to the tasks the students were required to do. The tasks (e.g., individual assignment, group assignment) did not necessarily require the students to use certain structures such as sentence-ending variables compared to other structures (e.g., clause conjunction) and hence they were unable to show improvement.

In terms of accuracy, the results from Study 1 suggest that in general, accuracy is one of the hardest measures to improve (i.e., it only showed significant
differences in the upper proficiency levels) and is one of the hardest writing features to measure, as seen in Study 2.

As seen from the above findings, developmental patterns varied depending on the writing features and measures used. The current project has theoretical, methodological and pedagogical implications, which are discussed in the following section.

10.2 Implications

10.2.1 Theoretical implications

Two different cognitive theories of SLA (i.e. the noticing hypothesis [Schmidt, 1990], and ‘practice’ in skill acquisition theory [DeKeyser, 2007]) were used to discuss the findings of Study 2. The study lends support to both theories. This suggests that awareness (i.e., noticing) via explicit instruction and reading, as well as written output practice is conducive to L2 writing development. The noticing hypothesis (Schmidt, 1990) and written output practice (DeKeyser, 2007) may explain why learners showed improvement particularly in terms of kanji complexity (e.g., KC RTTR), fluency and structure.

Through activities such as the kanji tests and readings, students were able to notice new vocabulary, then retrieve that new vocabulary and practice it when writing their assignments. This explains the observed improvement in syntactic complexity and fluency. The results also suggest that complex sentences were ‘noticed’ in the readings and through explicit instruction. Explicit teaching and examples of how to achieve coherence in class made students conscious about how to structure an essay. Thus, the findings support the claim that conscious learning plays an important part in writing development.
At the same time the lack of feedback, usually another means to enable learners to notice the gap or mismatch between the TL input and their output (Bitchener & Storch, 2016), explains why the learners showed no improvement in accuracy in the long term. Many L2 writing scholars (e.g., Bitchener & Storch, 2016; Leki, 1990) argue that feedback is a critical element for the development of written accuracy because it allows students to be aware (conscious) of their errors (noticing). Without regular feedback, it is difficult for learners to ‘notice the gap’.

Even though students in the current study had the opportunity to practice what they had learnt in their written assignments, classroom observations and student interview results indicated that the limited amount of writing time in class meant that the amount of practice may not have been sufficient. According to De Keyser (2007), the volume of engagement in writing practice plays a vital role in L2 language development. It is believed that practice will eventually result in automatisation of knowledge. Perhaps if students had more opportunities to practice their writing, they could have achieved even greater improvement.

10.2.2 Methodological implications

The current study has three methodological implications, particularly in terms of the measures used to assess L2 writing development. First, the results from Study 1 showed that various measures in syntactic complexity and accuracy only had significant differences in the lower proficiency levels. This suggests a need for measures that are able to distinguish writing development at higher-proficiency levels, and this is especially so for Japanese-specific measures.

Second, the findings from Study 2 showed that some measures improved only in the short term. This suggests that when using CAF measures to assess writing development over time, some measures are more informative (e.g., number of words,
number of sentences in fluency) than others (e.g., ratio of expressive endings, ratio of te-form in syntactic complexity). Thus, the current project suggests that CAF measures need to be used with caution, as some measures do not exhibit developmental changes in a linear way.

The longitudinal design adopted in Study 2 was an authentic classroom-based study. The majority of the previous studies (e.g., Storch & Tapper, 2009; Xudong, et al., 2010) involved a specific course of instruction (e.g., EAP course). Thus, the research design of the current study allowed us to understand what features of L2 Japanese writing develop in an authentic classroom situation.

10.2.3 Pedagogical implications

Four different teaching approaches were described in the literature review (Chapter 2). The teaching approach observed in the current study predominantly incorporated the product approach, with occasional elements of the genre approach (e.g., teaching learners in Japanese 3 the features of tourism texts and the assessment tasks; see Appendix I). The product approach was characterised by classroom practices such as weekly grammar instruction and workbook exercises (e.g., substitution exercises).

The results revealed that activities such as ERs and assignments focusing on particular genres might assist in learners’ writing development. Through these activities, students were able to acquire vocabulary, as well as learn how to structure sentences in Japanese. However, the findings suggest that writing development depends on sustained practice and feedback. For example, in Japanese 3 many activities enabled students to practice writing a persuasive piece of writing, such as a script for an oral assessment task and an individual assignment that focused on making a webpage of a tourist destination in Japan. Improvement in persuasiveness
was therefore seen in the short term (see more details in see Chapter 9, Section 9.1.5.2). However, further improvement in the long term was not evident because the activities in Japanese 4 did not give students opportunities to practice persuasive writing (see Chapter 9, Section 9.1.5.2)

As mentioned in the interviews, there was limited teacher feedback in both Japanese 3 and Japanese 4; yet feedback has been identified as important in the literature on L2 writing, as well as by the student (and teacher) participants. In classrooms with limited time for teachers to give feedback to students, peer feedback could be considered as an option. Researchers (e.g., Berg, 1999; Min 2005, 2006; Rahimi, 2013; Zhu 1995, 2001) have suggested that trained peer responses lead to a positive outcome in students’ revision type and quality. In the current study, there was an opportunity for students to provide peer feedback in Japanese 4, but there was insufficient peer feedback training (see Chapter 8, Section 8.3.4.5) to allow students to benefit from this process. As reported in the interviews, some students did not trust their peers’ comments. Thus, for peer feedback to be implemented effectively in class, it is suggested that students undertake peer feedback training beforehand. Another option to compensate for the limited teacher feedback could be discussing the common errors and providing examples in class, a suggestion made by Tanaka-sensei.

Teaching writing at an early stage is considered important for character-based languages (Cook & Bassetti, 2005). It may take more time to develop writing compared to alphabet-based languages, as character-based languages are considered more complex (Marriott et al., 1993). There was little emphasis on writing in the current study (Study 2), as noted in the observations and as commented by Hayashi-
sensei. Time allocated to writing and the type of writing activities in class are clearly challenges that need to be considered by the teachers and course designers.

10.3 Limitations of the Study

Inevitably, as with any research, the current study had several limitations. First, the writing task in Study 1 (used as a placement test at the university) may have been too simple (writing an email to a Japanese friend about a trip to Japan) to elicit sophisticated writing from students at higher-proficiency levels.

Second, only one topic was used for the pre- and post-test task. Using the same topic was needed to compare the pre-and post-test task. The pre- and post-test task was similar to what the students were learning in Japanese 3. Thus, through practice students were able to perform well in their post-test. Students may have performed differently if a different topic was used.

Third, because interviews with only two teachers were conducted in Japanese 3 (Semester 1) and one teacher in Japanese 4 (Semester 2), the interview data might lack generalisability and may possess a limited perspective.

Investigation on writing development for character-based learners and non-character-based learners was only conducted for one semester, as the number of participants in Japanese 4 was too small for division into two groups and comparisons. In order to examine the differences in writing development between the two groups, it would have been ideal for the study to be conducted for more than one semester with a larger sample size. Further investigation on the impact background has on motivation for writing may also have been useful. A learner’s language background may have an influence on motivation to write and affect writing development.
In the current study, limited numbers of measures (e.g., CAF) were used to examine writing development. These measures may not have been sufficient to examine writing development, as it may have masked development in other aspects of writing. In addition, the factors that contributed to development were not directly linked to individual students who improved their writing.

Even though assignments were collected in both Japanese 3 and Japanese 4, they were not used in the data analysis because in both subjects there was only one assignment each. Therefore, the assignments did not enable examination of learners’ writing development. In addition, assignments in Japanese 3 and Japanese 4 could not be compared for writing development because the topic was different.

Finally, a tape or video recording when students with working on their group assignments would have been useful, as it would have provided more data on the type and quality of peer feedback given. It was not possible, however, to have a tape recording of group work, as some students in the groups were not participants of the research project.

10.4 Suggestion for Further Research

Based on the findings of this research project, several suggestions can be made for future research. The majority of the studies in the past on L2 writing development have been in an ESL/EFL context with fewer studies beyond these contexts, especially in character-based languages like Japanese and Chinese. Due to the differences in languages, writing development can differ; therefore, more studies on L2 Japanese writing should be undertaken to shed light on how to measure L2 writing development in character-based languages.

Japanese is one of the most widely studied languages in Australia. Yet to date, there have been only a few longitudinal studies undertaken in authentic
Japanese L2 classroom settings. Future research building on this study and conducted in such settings could provide Japanese L2 teachers and researchers with a better understanding of teaching practices and activities that can lead to L2 writing development in their classes. The longitudinal study in this research project only investigated the writing development of lower-intermediate learners. Further research could be extended to learners with different proficiency levels.

In the current research project, only a cross-sectional and longitudinal design was used to explore writing development. Future research could use a case study approach to examine individual writing development using both quantitative and qualitative analysis. Such an approach would allow individual writing development to be examined in more detail.

Differences in motivation for writing between the character-based and non-character-based learners could be another area for future exploration. As character-based learners have an advantage over non-character-based learners, the motivation to write may differ between the two learners. This may affect writing development.

10.5 Concluding Statement

The current research project provides an empirical response to L2 writing scholars’ calls (e.g., Cimasko & Reichelt, 2011; Reichelt, 2001) for research on L2 writing in LOTE. This research project is among the few studies to investigate L2 Japanese writing development in a FL context using both cross-sectional and longitudinal mixed-method designs. Nevertheless, more research on L2 Japanese writing is needed, with attention paid to ecologically valid study designs; that is, longitudinal studies conducted in authentic classroom settings. The findings have a number of key methodological and pedagogical implications. It is the intention of
this research project to provide a number of pathways for future research to enhance our understanding of L2 writing development.
References


Coscom Language Service (2013). Retrieved from:


342


Ishida, K. (2007). Developing understanding of how the desu/masu and plain forms express one’s stance. In D. R. Yoshimi & H. Wang (Eds.), *Selected papers from the conference on pragmatics in the CJK classroom: The state of the art*


Moursssi, A. (2013). Theoretical and practical linguistic shifting from product/guided writing to process writing and recently to the innovated writing process.
approach in teaching writing for second/foreign language learners.


Nihongoryoku tesuto no jikkou wo tooshite, Chubu Daikaku Kyouiku Kenkyu,
7, 51–56.

Yamashita, J. (2013). Effects of extensive reading on reading attitudes in a foreign
language. Reading in a Foreign Language, 25(2), 248–263.

feedback in a Chinese EFL writing class. Journal of Second Language

and affect on willingness to communicate and second language
communication. Language Learning, 54(1), 119–152.

writers’ genre awareness, linguistic knowledge, and writing competence.

RELJ Journal, 38(1), 38–52.

Zamel, V. (1983). The composing processes of advanced ESL students: Six case
studies. TESOL Quarterly, 17(2), 165–188.


Zhao, H. (2010). Investigating learners’ use and understanding of peer and teacher
feedback on writing: A comparative study in a Chinese English writing

English majors’ writing. English Language Teaching, 8(3), 89.


Appendices

Appendix A

Writing

This task allows you to demonstrate how well you can write in Japanese, including the range of vocabulary and structures you are able to use accurately.

Write an email based on the following information:

You are planning to go on a trip to Japan. You have a Japanese-speaking friend in Japan who might be able to help you plan your trip. Write an email in Japanese to this friend outlining your plan (explaining what you want to do, why, where and how) and asking for advice and information on certain points.

If you like, you can also include:

- consideration of Japanese culture (e.g. gift giving when visiting someone)
- recent social topics (e.g. earthquake, strong Yen, political leadership change)

You have a total of 30 minutes to complete this task. Please write as much as possible within this time limit. Write your answer in the space provided below.

__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
Appendix B

Background Questionnaire

Please answer the following questions: Tick ✓ where appropriate and tick as many options applicable.

A. PERSONAL DETAILS

Family name: _____________________  First given name: ______________

Country of birth: ___________________  Student ID number: _____________

Unimelb email:

Gender: □ Male  Residency status: □ Domestic

□ Female  □ International

Course type: □ Undergraduate  Enrolment status: □ Full-time

□ Graduate  □ Part-time

□ Community access

Faculty: □ Architecture, Building & Planning

□ Arts

□ Business & Economics

□ Educatio

□ Engineering

□ Information Technology

□ Land & Environment

□ Law

□ Medicine, dentistry & health sciences

□ School of graduate research

□ Science

□ Veterinary science
Degree/Diploma: _____________________

Are you taking Japanese as a breadth subject? □ Yes □ No □ Unsure

Are you intending to major in Japanese? □ Yes □ No □ Unsure

Subject: □ Japanese 1/2  
□ Japanese 3/4  
□ Japanese 5/6  
□ Japanese 7/8

B. LANGUAGE BACKGROUND
1. What languages do you speak? _________________________________________
2. What is your first (native) language? ________________________________

3. Do any members of your family speak Japanese?
   □ No → go to question 6
   □ Yes
      □ Brother/sister(s)
      □ Parent(s)
      □ Grandparent(s)
      □ Other ___________

4. Do you speak Japanese with members of your family?
   □ No → go to question 6
   □ Yes
      □ Brother/sister(s)
      □ Parent(s)
      □ Grandparent(s)
      □ Other ___________

5. What type of Japanese do you speak with members of your family?
   □ A regional Japanese dialect
6. Have you ever lived in Japan?

☐ No → go to question 7
☐ Yes
  ☐ for less than 3 months
  ☐ for 3 months–1 year
  ☐ for 1-2 years
  ☐ for 3-4 years
  ☐ for 5-6 years
  ☐ for more than 6 years

C. EDUCATION

7. Have you ever studied Japanese in Japan?

☐ No → go to question 9
☐ Yes (in total)
  ☐ for less than 3 months
  ☐ for 3 months–1 year
  ☐ for 1-2 years
  ☐ for 3-4 years
  ☐ for 5-6 years
  ☐ for more than 6 years

8. What type(s) of school did you attend in Japan?

☐ primary school
☐ secondary school
☐ university
☐ private language school
☐ other ____________________

9. Have you ever studied Japanese in a non-Japanese-speaking country?
10. For how long did you study in a non-Japanese-speaking country? (in total)
   - for less than 3 months
   - for 3 months - 1 year
   - for 1 - 2 years
   - for 3 - 4 years
   - for 5 - 6 years
   - for more than 6 years

11. At what type(s) of school have you studied Japanese?
   - primary school
   - secondary school
   - university
   - Saturday school
   - Other private language school
   - Private individual tuition
   - other ____________________

12. Have you completed Year 12 Japanese in an Australian school?
   - No → go to question 13
   - Yes Year ______

   □ VCE know
      Score: ______  □ raw or □ scaled or □ don’t
   □ NTCE          Score: ______
   □ Senior Certificate Score: ______
   □ IB            Score: ______
   □ HSC           Score: ______
   □ Year 12 certificate Score: ______
13. Have you ever been taught subjects other than Japanese in the Japanese language (e.g. maths lessons with Japanese as the medium instruction)?

☐ No → questionnaire finished

☐ Yes

☑ primary school
☐ secondary school

14. For how long were you taught other subjects in Japanese? (in total)

☐ for less than 3 months
☐ for 3 months-1 year
☐ for 1-2 years
☐ for 3-4 years
☐ for 5-6 years
☐ for more than 6 years
Appendix C

If August 2013

Dr. H. Storch
Institute of Linguistics
The University of Melbourne

Dear Dr. Storch,

I am pleased to advise that the Humanities and Social Sciences Human Ethics Sub-Committee approved the following project:

Project Title: Writing Development of second language Japanese learners in Australia

Project Leader: Dr. H. Storch, Dr. V. Winch

Date of Approval: 14 August 2013

The Project has been approved for the period: 31-Aug-2013 to 31-Dec-2013

It is your responsibility to ensure that all people associated with the Project are made aware of what has been approved.

Research projects are normally approved to 31 December of the year of approval. Projects may be renewed yearly subject to a local review of the research and receipt of a satisfactory annual report. The project is continued beyond five years a new application will normally need to be submitted.

Please note that the following conditions apply to your approval. Failure to comply with these conditions may result in suspension of reimbursement of approval and/or termination of approval.

(a) Limit of Approval: Approval is limited strictly to the research as submitted in your Project application.

(b) Variation to Project: Any subsequent variations or modifications you may wish to make to the Project must be notified. Normally the Human Ethics Sub-Committee will review your request and the project will be continued.

(c) Review of Ethics: Researchers must report immediately to the Project Committee any adverse events or serious problems which may affect the ethical acceptance of the research or the safety or rights of participants. This may result in suspension of the Project or intervention by the Project Committee.

(d) Monitoring: All projects are subject to monitoring at any time by the Human Research Ethics Committee.

(e) Annual Reports: Please be aware that the Human Research Ethics Committee requires that researchers submit an annual report on each of their projects at the end of the year. The conclusion of a project 6 continues for less than 6 months. Failure to submit an annual report will mean that ethics approval will lapse.

(f) Auditing: All projects may be subject to audit by members of the Sub-Committee.

If you have any queries on these matters, or require additional information, please contact us using the details below.

Please quote the reference number and the title of the Project in any future correspondence.

On behalf of the Sub-Committee I wish you well in your research.

Yours sincerely,

[Signature]

Humanities and Social Sciences HESC
[Signature]

Professor [Name]
[Email: [email protected]]

[Telephone]

Officer for Research Ethics and Integrity

[Signature]

[Telephone]

[Email: [email protected]]

[Telephone]

[Email: [email protected]]
Appendix D

Guidelines for coding T-units (Adapted from Cumming et al., 2006 with some modification for Japanese)

a. A T-unit equals a main clause with all its subordinate clauses.

Example: My school, which I liked very much was in Tokyo.

Watashi ga daisuki datta gakko wa Tokyo ni arimashita.

1 T-unit = 1 T
1 clause/ 1 clause (continuation of the first clause) = 2 C

b. Run-on sentences and comma splices have as many T-units as there are independent clauses.

Example: My school was in Saudi Arabia, I liked it very much.

Watashi no gakko wa Tokyo ni arimashita, totemo sukideshita.

1 T-unit/ 1 T-unit = 2 T
1 clause/ 1 clause = 2 C

c. Follow these rules for sentence fragments:

i. If the verb or copula is missing, count the sentence as one T-unit (and, of course as one clause).

Example: My school the best in Tokyo.

Watashi no gakko Tokyo ichiban.

ii. If a noun phrase (NP) is standing alone, attach it to the preceding or following T-unit as appropriate.

Example: My school in Saudi Arabia, I like it very much.

Tokyo no watashi no gakko, totemo sukidesu.
iii. If a subordinate clause is standing alone, attach it to the preceding or following sentence but count it as a separate clause.

0 T-unit → T-unit
1 clause / 1 clause = 2C

d. When there is a grammatical subject deletion in a coordinate clause, count the entire sentence but count each clause separately.

Example: First we went to our school and then went out with our friends.
最初に私たちの学校に行ってから友達と出かけました。

Saisho ni watashi tachi no gakko ni itekara tomodachi to
dekakemashita.
1 T-unit
1 clause / 1 clause = 2C

e. Count so and but as coordinating conjunctions, and so that as subordinating conjunction unless so is obviously meant.

Example: We go to school so that we can learn. (so that = so that)
わたしたちは学ぶために学校に行きます。

Watashi tachi wa manabu tameni gakko ni ikimasu.
Note: so that (‘tameni’) is underlined in the Japanese sentence.
1 T-unit
1 clause / 1 clause = 2C

f. Do not count tag-questions as separate T-units or separate clauses.

Example: We go to school so in Tokyo, don’t we?
私たちは東京の学校に行きます、ですよね。

Watashi tachi wa Tokyo no gakko ni ikimasu, desuyone.
Note: In Japanese “?” is not used at the end of the sentence for questions.
1 T-unit
1 clause
g. Do not count tag-statements as separate T-units or separate clauses.

**Example:** We go to school so in Tokyo, and Mary does, too.

私たちは東京の学校に行きまます、メリーもです。

*Watashi tachi wa Tokyo no gakko ni ikimasu, Meri mo desu.*

1 T-unit

1 Clause

h. Count direct quotes as T-units (and, of course, as separate clauses).

**Example:** John said, “I am hungry”

ジョンは、「お腹が空いている」と言いました。

*John wa “Onaka ga suiteiru” to iimashita.*

1 T-unit/ 1 T-unit
Appendix E

Coding of *kanji* errors adapted from Kubota (2005)

<table>
<thead>
<tr>
<th>Types of <em>kanji</em> errors</th>
<th>Example of an error</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confusion with morphographically similar <em>kanji</em></td>
<td>1. 恋わりました</td>
<td>恋: the <em>kanji</em> on the bottom is 心 whereas the correct <em>kanji</em> should be 变</td>
</tr>
<tr>
<td></td>
<td>2. 変わりました</td>
<td></td>
</tr>
<tr>
<td>A misshapen component of the <em>kanji</em></td>
<td>1. 鳳物</td>
<td>When learners handwrite <em>kanji</em>, it is sometimes poorly written</td>
</tr>
<tr>
<td></td>
<td>2. 建物</td>
<td></td>
</tr>
<tr>
<td>Problems with okurigana</td>
<td>1. 入いる</td>
<td>When a word is expressed with <em>kanji</em> and hiragana, hiragana functions as okurigana, an ending that supplements the word. In this case, there is an additional hiragana in the word, 入いる instead of 入る</td>
</tr>
<tr>
<td></td>
<td>2. 入る</td>
<td></td>
</tr>
<tr>
<td>Transfer from Chinese characters to Japanese</td>
<td>1. 傳統的</td>
<td>Background learners sometimes get confused with the Chinese characters and <em>kanji</em></td>
</tr>
<tr>
<td></td>
<td>2. 伝統的</td>
<td>*please note: this will be especially relevant for research question 3, which compares learners with different backgrounds.</td>
</tr>
<tr>
<td>Homonyms</td>
<td>1. 始めて</td>
<td>始めて和 初めて are both pronounced ‘hajimete’. 始めて means ‘to begin’ but 初めて means ‘first time’.</td>
</tr>
<tr>
<td></td>
<td>2. 初めて</td>
<td></td>
</tr>
<tr>
<td>The absence of one <em>kanji</em> from a compound KW</td>
<td>1. 小校</td>
<td>The <em>kanji</em> 学 is missing in the 小校</td>
</tr>
<tr>
<td></td>
<td>2. 小学校</td>
<td></td>
</tr>
</tbody>
</table>
### Coding of kana errors adapted from Kubota (2005)

<table>
<thead>
<tr>
<th>Types of kana errors</th>
<th>Example of an error</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. error</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. correct version</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A misshapen kana</td>
<td>1. 変わりました</td>
<td>When learners handwrite kanji, it is sometimes poorly written</td>
</tr>
<tr>
<td></td>
<td>2. 変わりました</td>
<td></td>
</tr>
<tr>
<td>A missing component</td>
<td>1. なります</td>
<td>The hiragana す is missing at the end.</td>
</tr>
<tr>
<td></td>
<td>2. なりました</td>
<td></td>
</tr>
<tr>
<td>A wrong kana</td>
<td>1. 思し</td>
<td>Hiragana し is used instead of い.</td>
</tr>
<tr>
<td></td>
<td>2. 思い</td>
<td></td>
</tr>
<tr>
<td>Transfer from Chinese characters to Japanese</td>
<td>1. 傳統的</td>
<td>Background learners sometimes get confused with the Chinese characters and kanji</td>
</tr>
<tr>
<td></td>
<td>2. 伝統的</td>
<td>*please note: this will be especially relevant for research question 3, which compares learners with different backgrounds.</td>
</tr>
<tr>
<td>Writing a loan word approximating the original word rather than using the Japanised version</td>
<td>1. ヴァラエティー</td>
<td>When the loan word ‘variety’ is written as ヴァラエティー, the ヴァ will be close to the English ‘va’, but the Japanised word is written with バ with the ‘ba’ pronunciation, hence ‘baraety’ in romaji.</td>
</tr>
<tr>
<td></td>
<td>2. バラエテイー</td>
<td></td>
</tr>
</tbody>
</table>
Appendix F

School of Languages and Linguistics

Consent form for persons participating in the following research project

PROJECT TITLE: Writing Development of L2 Japanese learners (tertiary level) in Australia

Name of participant: 

Name of investigator(s): Dr Neomy Storch, Dr Ute Knoch, and Yuka Kikuchi

1. I consent to participate in this project, the details of which have been explained to me, and I have been provided with a written plain language statement to keep.

2. I understand that after I sign and return this consent form it will be retained by the researcher.

3. I understand that my participation will involve a 30 minute writing task in the beginning and at the end of the semester, permitting the researchers to collect copies of my graded work completed throughout the semester, completing a background questionnaire and participating in an interview.

4. I acknowledge that:
   (a) the possible effects of participating in the study have been explained to my satisfaction;
   (b) I have been informed that I am free to withdraw from the project at any time without explanation or prejudice and to withdraw any unprocessed data I have provided;
   (c) the project is for the purpose of research;
   (d) I have been informed that the confidentiality of the information I provide will be safeguarded subject to any legal requirements;
   (e) I have been informed that with my consent the interview will be audio-taped and I understand that audio-tapes will be stored at the University of Melbourne and will be destroyed after five years;
   (f) I have been informed that with my consent the writing scripts collected will be stored at the University of Melbourne and will be destroyed after five years.
   (f) my name will be referred to by a pseudonym in any publications arising from the research;
   (g) I have been informed that a copy of the research findings will be forwarded to me, should I agree to this.

I consent to this interview being audio-taped □ yes □ no

I wish to receive a copy of the summary project report on research findings □ yes □ no

Participant signature: 

Date: 
Appendix G

School of Languages and Linguistics

Consent form for persons participating in the following research project

PROJECT TITLE: Writing Development of L2 Japanese learners (tertiary level) in Australia

Name of participant: 

Name of investigator(s): Dr Neomy Storch, Dr Ute Knoch, and Yuka Kikuchi

1. I consent to participate in this project, the details of which have been explained to me, and I have been provided with a written plain language statement to keep.

2. I understand that after I sign and return this consent form it will be retained by the researcher.

3. I understand that my participation will involve distributing task 1 in the beginning of the semester and be involved in an interview at the end of the semester, allow the researcher to copy students’ writing assessment (with my comments) and I agree for the researcher to observe my classes.

4. I acknowledge that:
   (a) the possible effects of participating in the study have been explained to my satisfaction;
   (b) I have been informed that I am free to withdraw from the project at any time without explanation or prejudice and to withdraw any unprocessed data I have provided;
   (c) the project is for the purpose of research;
   (d) I have been informed that the confidentiality of the information I provide will be safeguarded subject to any legal requirements;
   (e) I have been informed that with my consent the interviews will be audio-taped and I understand that audio-tapes and the observation notes will be stored at the University of Melbourne and will be destroyed after five years;
   (f) my name will be referred to by a pseudonym in any publications arising from the research;
   (g) I have been informed that a copy of the research findings will be forwarded to me, should I agree to this.

I consent to this interviews being audio-taped □ yes □ no (please tick)

I wish to receive a copy of the summary project report on research findings □ yes □ no (please tick)

Participant signature: 

Date:
Appendix H

School of Languages and Linguistics

Consent form for persons participating in the following research project

PROJECT TITLE: Writing Development of L2 Japanese learners (tertiary level) in Australia

Name of participant: __________________________

Name of investigator(s): Dr Neomy Storch, Dr Ute Knoch, and Yuka Kikuchi

1. I consent to participate in this project, the details of which have been explained to me, and I have been provided with a written plain language statement to keep.

2. I understand that after I sign and return this consent form it will be retained by the researcher.

3. I understand that my participation will involve a 30 minute writing task in the beginning of the semester, and two 30 minute writing tasks at the end of the semester, permitting the researchers to collect copies of my graded work completed throughout the semester, completing a background questionnaire and participating in an interview.

4. I acknowledge that:

   (a) the possible effects of participating in the study have been explained to my satisfaction;
   
   (b) I have been informed that I am free to withdraw from the project at any time without explanation or prejudice and to withdraw any unprocessed data I have provided;
   
   (c) the project is for the purpose of research;
   
   (d) I have been informed that the confidentiality of the information I provide will be safeguarded subject to any legal requirements;
   
   (e) I have been informed that with my consent the interview will be audio-taped and I understand that audio-tapes will be stored at the University of Melbourne and will be destroyed after five years;
   
   (f) I have been informed that with my consent the writing scripts collected will be stored at the University of Melbourne and will be destroyed after five years.
   
   (f) my name will be referred to by a pseudonym in any publications arising from the research;
   
   (g) I have been informed that a copy of the research findings will be forwarded to me, should I agree to this.

I consent to this interview being audio-taped □ yes □ no (please tick)

I wish to receive a copy of the summary project report on research findings □ yes □ no (please tick)

Participant signature: __________________________  Date: __________________________
## Appendix I

### Japanese 3 Course outline

**Draft: 28/02/2013**

**Japanese 3 (JAPN1007/JAPN2007)**

**Teaching Schedule (may be updated without notice)**

<table>
<thead>
<tr>
<th>Week</th>
<th>Topic</th>
<th>Lecture</th>
<th>Tutorial 1</th>
<th>Tutorial 2</th>
<th>Assessments Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Legends and folktales (げんき 第1 3 閣)</td>
<td>Reading tourist brochure How to study kanji &amp; survive</td>
<td>Self-introduction 動词接続形 (p.30) ～し～ (p.31) ～そうですね (p.32) ～てみる (p.34) 表現ノート ギターを弾く、上手い (p.36)</td>
<td>Self-introduction 阅读旅游指南 新词接续形 (p.30) ～し～ (p.31) ～そうですね (p.32) ～てみる (p.34) 表現ノート ギターを弾く、上手い (p.36)</td>
<td>4/3**</td>
</tr>
<tr>
<td>2</td>
<td>Amaterasu Okuninushi &amp; Susanoo (げんき 第1 3 閣)</td>
<td>About history of Tenno (Emperor) Historical stories in brochures</td>
<td>Placement test field trial</td>
<td>Map of train lines 阅读关于运输在日本 ～なら (p.34) ～週間に三回 (p.35)</td>
<td>11/3**</td>
</tr>
<tr>
<td>3</td>
<td>Yamato-no orochi (げんき 第1 4 閣)</td>
<td>More historical details in brochures</td>
<td>はしい (p.54) ～かもしれませんが (p.56) あげる・くれる・もらう (p.56)</td>
<td>Giving and Receiving - culture and practice 基本历史的日本</td>
<td>18/3**</td>
</tr>
<tr>
<td>4</td>
<td>Shinkansen (げんき 第1 4 閣)</td>
<td>History and development of Shinkansen</td>
<td>たらどうですか (p.58) 車が三回もいます 車が三回しかいません (p.58)</td>
<td>Good Friday: NO CLASS HW: Reading about someone’s issues and offer suggestions</td>
<td>25/3** Easter break Mon: CP1</td>
</tr>
<tr>
<td>5</td>
<td>Train timetable (げんき 第1 5 閣)</td>
<td>Train timetable Dealing with unknown kanji</td>
<td>食べよう・行こう (p.73) ～と思ってています (p.73) ～ておく (80)</td>
<td>Map of Japan, train timetable 阅读关于不同的地方 Explaining places you like</td>
<td>11/4**</td>
</tr>
<tr>
<td>6</td>
<td>Shrines and temples (げんき 第1 6 閣)</td>
<td>Shrines and temples: contrasting Shinto and Buddhism</td>
<td>表現ノート casual speech (p.82)</td>
<td>Reading tourist information web page in other cultures - finding out genre</td>
<td>15/4** Mon: Speaking 1</td>
</tr>
</tbody>
</table>
# Appendix J

## Japanese 4 course outline

Updated 21/08/2013

**Japanese 4 (JAPN10008/JAPN20008)**

**Teaching Schedule (may be updated without notice)**

<table>
<thead>
<tr>
<th>Week</th>
<th>Topic</th>
<th>Lecture</th>
<th>Tutorial 1</th>
<th>Tutorial 2</th>
<th>Assessments Due</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Week 1:</strong> (げんき 第 1 週)</td>
<td>Introduction to Japanese 4 topics Research about Japanese writing system, topics are introduced each week in lecture</td>
<td>Hononif verbs (p.168) ～てくれてありがとう (p.171)</td>
<td>Reading: A letter to host family - Formal letters in Japanese Aspects of Japanese writing</td>
<td>Make groups of 3 with students from another culture</td>
<td></td>
</tr>
<tr>
<td><strong>Week 2:</strong> (げんき 第 1 週)</td>
<td>How kanji became a part of Japanese; Can we abolish kanji?</td>
<td>Giving respectful advice お〜ください (p.170) ～てよかったです (I am glad that) (p.171) ～はずですか (p.172)</td>
<td>Reading: An email to a new acquaintance</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week 3:</strong> (げんき 第 2 週)</td>
<td>Development of katakana and hiragana</td>
<td>Extra-modest expressions, humble expressions (p.186-191)</td>
<td>Writing a formal letter</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week 4:</strong> (げんき 第 2 週)</td>
<td>About Keigo About Role play assignment 〜ないで (p.191) Question Clause か〜かどうか verb (p.192) Visiting someone's house</td>
<td>Feeding: Rakugo: Neko no cara</td>
<td>Mon: CPI</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week 5:</strong> (げんき 第 2 週)</td>
<td>Studying kanji</td>
<td>Noun という item (p.193) ～やすい・にくい (p.193)</td>
<td>Viewing: Rakugo: Neko no chawon Writing practice</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week 6:</strong> (げんき 第 2 週)</td>
<td>Onomatopoeia in manga Passive (p.212) ～てある (214) ～間に (p.216)</td>
<td>Reading: Yakusoku</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week 7:</strong> (げんき 第 2 週)</td>
<td>Japanese characters in computing</td>
<td>Adj する (p.216) ～てほしい (p.217)</td>
<td>Reading: Computer and Japanese</td>
<td>Mon: CP2</td>
<td></td>
</tr>
<tr>
<td><strong>Week 8:</strong> (げんき 第 2 週)</td>
<td>Kotodama and Mojibana</td>
<td>Casusative (p.224) Verb 使わない 〜にさい (p.236) 〜ば (p.236)</td>
<td>Reading: Tomomi's diary</td>
<td>Fri: Speaking 1</td>
<td></td>
</tr>
</tbody>
</table>
Appendix K

Writing task

Name: ________________________  Student ID: _______________

Write a promotional newspaper/magazine article about your hometown or a place you know very well. The target audiences of the article are tourists coming to the region for a visit.

You may want to include information such as:

- The name of the town/attraction
- Detailed information on the town/attraction
- What is there to do
- How to get to the town/attraction
- How much it may cost to do various activities

____________________________________________________________________
____________________________________________________________________
____________________________________________________________________
____________________________________________________________________
____________________________________________________________________
____________________________________________________________________
____________________________________________________________________
____________________________________________________________________
____________________________________________________________________
____________________________________________________________________
____________________________________________________________________
____________________________________________________________________
____________________________________________________________________

386
Appendix L

Semester 1 (Japanese 3) Student Participant Interview Questions

Group assignments:

1) During this semester you had three group assignments. By working as a group, do you think it helped you improve your writing? If so what areas did it help you improve your writing?

2) I believe that you had to use Japanese websites in order to write up your group assignments. By reading Japanese websites, did it help you write better? If so how? Or was it difficult to use Japanese websites? and if so what were the difficulties that you faced?

Individual Assignment:

Your last assignment was an individual assignment (designing a web page, introducing a tourist destination).

- Compared to writing as a group was it easier on your own or difficult? Why?
- Were you able to use the writing skills that you developed when you did your group assignment? If so what were the skills that you think you have acquired during the semester?

Writing Task:

I am not sure how much you remember what you wrote in week 1 for your writing task but do you think you were able to write better the second time?

- I think you wrote better. What do you think helped you?
- Which areas do you think you felt you improved the most? If you feel you didn’t improve, why?

Do you think if you wrote the writing task after the exam you will be able to write even better?

Writing practice during class:

- Do you think the writing practices that you did in class (reflection of class) helped you improve your writing?
- If yes, how?
- If not, why? Could you suggest some writing activities that may be introduced in class in the future?

**Kanji**

I understand that you had to learn 8 – 10 kanji characters every week in seminar 1. Usually how long does it take for you to learn all those kanji characters?

- For non-character-based: Is memorising kanji a struggle for you? Or do you enjoy it? And why?

- For character-based: How do you go about remembering these kanji? What difficulties do you have with these characters?

**Writing by hand and typing**

Do you find typing in Japanese is easier than writing by hand in Japanese (e.g. you can choose kanji from the computer?)? Why and why not?

**L1 Writing Expertise**

Do you like to write in your native language?

- **YES answer:**
  

- **NO answer:**
  
  Why don’t you like writing? Do you also struggle to write in Japanese? If so what aspects do you struggle with and why?

**General question:**

- What were the class activities or assignments you found useful in improving your writing? Why?

- What are the areas (e.g. kanji, hiragana, structure, content) that you feel you have improved in over the semester?

- What kind of writing practice did you incorporate in your learning to help you improve?

- Are there any areas in writing that you would like to improve further? If so, what are they and what kind of class activities or assignments do you think may help you improve?
Outside class:

- Japanese animation/drama: (For those of who watch them)
  I believe you often watch Japanese animation/drama. Although it is not directly related to writing, do you think it helps to improve your writing in Japanese? If so, how/why?

- Friends:
  Do you have any Japanese friends you often have contact with? If so, do you write to them? Does this help you improve writing in Japanese and how?
Appendix M

Semester 2 (Japanese 4) Student Participant Interview Questions

Writing practice in class:
1) Did you find the writing practices (3 sentences - reflection of class) during class useful in improving your writing? If so why? If not, why not?
2) Last semester you had the same writing practices but this semester you had to do a peer feedback on those writing practices. What kind of feedback were you able to give to the other person? By correcting and being corrected, do you think it helped you with your writing? If so how?

Group assignments:
1) The group assignment was slightly different from last semester. Do you think the research based assignment helped you improve your writing? If so how?
2) Did you mainly use English based resources or Japanese based resources for your assignment? Did the resources that you used help you with your writing in Japanese? If so how?
3) The first assignment and the second assignment were to prepare for your individual assignment. Did they help you with writing up your individual assignment? If so how?
4) Oral video: You had to make a 5 minute oral video again this semester. Although the assessment itself was a speaking assessment, by writing up the script, did it help you with improving your writing? Why and why not? Did you find anything different from last semester?
4) Did working as a group help you with your writing? Why and why not?

Individual Assignment:

Peer feedback:
1) You had to give peer feedback on two individual assignments. Was it difficult or easy to give feedback? What were the difficulties that you faced?
2) I believe that the chosen topics were different for each individual which means that you had to provide feedback on two topics that may have been not familiar to you. Did that make any difference to the way you provided feedback? If so how?

3) By giving feedback do you think it also helped you improve your writing? If so how?

4) Did the feedback that you received enabled you to write better for your final assignment? If so how? If not, why not?

**Writing Task:**

1) Having studied Japanese 3 and Japanese 4, do you think you were able to write better this semester compared to last semester? If so which areas do you think you felt you improved the most? What made you improve?

2) I have given you two different topics – one in semester one, another in semester two. Did the topics influence the way in which you wrote? If so how? If not, why not?

**Lecture:**

The lectures in Japanese 4 were mostly about Japanese writing system. By learning about different Japanese writing system, did it help you with your writing? If so how? If not, why not?

**Kanji**

I understand that you had to learn 8 – 10 kanji characters again every week this semester in seminar 1. By building up the number of kanji characters that you have learnt over two semesters, did it help you improve your writing? If so how?

**General question:**

1) What were the class activities or assignments you found useful in improving your writing? Why?

2) What are the areas (e.g. kanji, hiragana, structure, content) that you feel you have improved in over the two semesters?

3) What kind of writing practice did you incorporate in your learning to help you improve?

4) Are there any areas in writing that you would like to improve further? If so, what
are they and what kind of class activities or assignments do you think may help you improve?

**Outside class:**

Do you engage in other activities outside class to improve your writing? If so, what are they?
Appendix N

Semester 1 (Japanese 3) Teacher Interview Questions

Teaching background information:
1) How many years have you taught Japanese?
2) How many years have you taught Japanese at the current university?
3) How many tutorials did you teach in Semester 1, 2014 (Japanese 3)?

Writing Development:
1) What influences writing development at this level (or what factors help students to develop their writing) and how/why
2) What aspects of writing do students develop over the course of a semester?
   - Do you think they are able to write more towards the end of the semester?
   - How about accuracy? If it is kanji?
   - How much emphasis is placed on writing skills in comparison to other aspects of language (speaking, grammar, vocabulary, reading, listening etc?). What do you think of that?
3) Do you think the teaching of Japanese writing is important at level 3 (Japanese 3), and if so why/why not?

Group assignments:
1) During semester 1, students had several group assignments. By having them work as a group do you think it helped them improve their writing? If so, in what ways?
2) Do you know how the different groups approach this activity?
3) I believe that students had to use Japanese websites in order to write up their group assignments. By reading Japanese websites, do you think that it helped them with their writing? If so, in what ways?
4) One of the speaking assessments was to make a video in a group. Although the assessment itself was “speaking”, since students had to write a script beforehand, do you think it helped them improve their writing as well? If so, how?
5) How do you think students approach this activity?
Individual assignment:
1) The individual assignment was to create a webpage. By creating a webpage in Japanese, what kind of writing skills were students able to develop?

What specific difficulties do you think students experience with this activity?

2) Do you think that students were able to use the writing skills that they may have acquired while doing the group assignment for their individual assignment? If so, what were they?

Class activities:
1) What aspects of the class work do you think supported or did not support students’ writing development?

2) What do you think was the main aim of the class reflection task (3 sentences each week at the end of the class)?

Do you think you will be using this task next time and why?

3) Do you think providing feedback was useful?

If so, what kind of feedback do you think was useful?

Students from various backgrounds:
1) How do you deal with students from various backgrounds?

2) Do you think some students struggle with kanji compared to others depending on their background?

3) In your teaching experience, how does a student’s language background affect their writing ability and their ability to be taught? And their ability to develop their writing ability?

Typing in Japanese and writing by hand
1) Do you find differences in students’ writing (e.g. errors) when they type and write by hand?

If so what are they?

2) As typing and writing by hand is different in Japanese, do you think students have specific problems with either way of writing?

3) Since we communicate using technology these days, how important do you think
being able to write by hand in Japanese for learners of Japanese? Why?

**Outside class:**
1) It seems that many students learn Japanese through Japanese anime and drama. Although it is not directly related to writing, do you think it helps improve their writing? If so, in what ways?

2) In your teaching experience, do you know what other kind of activities students do to improve their writing?

3) What kind of learning materials outside class have you recommended in the past (if you have) so that students could improve their writing?

**Motivation:**
In your teaching experience, what kind of class activities or assignments do you think motivates students to write in Japanese?
Appendix O

Semester 2 (Japanese 4) Teacher Interview Questions

Assignments:

Group assignment:
- What is your view on group assignment?
- Why is it used for both Japanese 3 and Japanese 4?
- What do you consider its advantages and disadvantages?

1) The first assignment (Semester 1) and second assignment (Semester 2) was done in a group. Although it was mostly done in English, they were to prepare for the individual assignment. By making them work as a group for Semester 1 and Semester 2, do you think it helped students write up their individual assignment? If so how?

Individual assignment:

1) The individual assignment was different in Japanese 3 and Japanese 4. What kind of writing skills were you hoping for students to acquire in Japanese 3 (web design) and Japanese 4 (research based assignment) ?

Peer feedback:

- What is your view on peer feedback?
- Why is it used in Japanese 4 but not in Japanese 3?
- What do you consider its advantages and disadvantages?

1) Last semester students had the same writing practices (3 sentences – reflection of class) but this semester in some weeks, they had to do a peer feedback. Do you think peer feedback helped them with their writing? If so how?

2) Do you think 3 sentences were enough for a peer feedback task?

3) This semester, each student had to give peer feedback on two individual assignments. I believe that most students did not have previous experience in giving feedback. In your opinion, do you think it was difficult or easy for them to give feedback? Why? Did you notice anything while you were reading students’ feedback?
4) I believe that students had to answer questions (e.g. Do you feel this paper relies on evidence, or on opinion or intuition?) to do their peer feedback task. Most of the questions are content based, why is that?

5) The peer feedback that students did in class (reflection of class – 3 sentences) was done face to face. However, the peer feedback on the individual assignments was anonymous. Do you think peer feedback done face to face and anonymous makes a difference in the way in which students provide feedback? If so how?

6) I believe that the chosen topics were different for each individual which means that students had to provide feedback on two topics that may not have been familiar. Do you think that made a difference in the way they provided feedback? If so how?

7) By students giving feedback to others, do you think it helped them improve their writing? If so how?

8) Do you think after the students received peer feedback, it enabled them to write better for their final assignment? If so how? If not, why not?

**Speaking Assessment:**
1) One of the speaking assessments was to make a video in a group. Although the assessment itself was “speaking”, since students had to write a script beforehand, do you think it helped them improve their writing as well? If so how?

**Lecture:**
The lectures in Japanese 4 were mostly about Japanese writing system. By learning about different Japanese writing system, do you think it helped students with their writing? If so how? If not, why not?

**Kanji:**
- What is your view on kanji learning? How important is it?
- I understand that students had to learn 8-10 kanji characters again every week this semester in seminar 1. By building up the number of kanji characters that they have learnt over two semesters, do you think it helped them improve their writing?

**General question:**
1) In your opinion, which class activities or assignments enabled students to improve their writing?
2) Usually which areas (e.g. *kanji*, *hiragana*, structure, content) do students improve by studying Japanese 3 and Japanese 4?

3) Do you know whether students incorporated any other writing practices outside class to improve their writing? If so what kind of writing practices did they do?

4) Are there any class activities or assignments that you would like to include in the future to enable students to improve their writing? If so what kind of class activities or assignments do you think are beneficial?
Appendix P

Follow-up Teacher Interview Questions (2016)

1) Did you design all the classroom activities and the assessment tasks for Japanese 3 and Japanese 4?

2) Have you been designing the classroom activities/assessment tasks ever since you started teaching at this university? If so, what areas (listening, writing, speaking, reading) have you been focusing when designing them?

3) What were the reasons for choosing specific assignments?

4) What were you hoping that the students will achieve from these assignments?

5) Why were most of the group assignments (2 out of 3) written in English not Japanese?

6) Did you use the same assignments last year (2015) or did you change them? How about the peer feedback assignment? Did you use it again last year?

If changed them – why did you changed them?

If not – why did you use the same assignments?

7) How about this year?

8) How about classroom activities? What was your focus on the design of the classroom activities?

9) Any changes that you made last year (2015)? (in terms of classroom activities)

10) How about this year (2016)?

12) How much emphasis do you put on writing in terms of designing classroom activities and assignments? Were there any specific skills you focused on?
Appendix Q

Coding examples for each rating scale.

Please note: The examples were taken directly from the learners’ writing script. Therefore some sentences are not necessarily written correctly. The English translation was translated as close as possible but some errors were unable to be translated.

Content: Sufficiency

Score (0): No response or just a few words/sentences

Malaysia is a very interesting country. Since the weather is always hot, there were many customers. Malaysia’s culture is very special. Chinese as well as Indians and Malay. It is full of shopping and delicious foods. It is cheap.

This writer had a score of 0 because she/he was only able to write a few sentences. Some sentences were also incomplete.
日高市は埼玉県の西にあります。人を6万人がいます。東京から電車で行ける。
日高市は山と神社とおいしい食べ物があります。日和田山は日高市のうめな山です。むりょうです。
高麗神社は日高市のうめな神社です。よく神社にまつりがあります。いいぱいおいしい物を食べられる。
日高の名物はくりと高麗鍋です。

Hidaka-shi is situated in the west of Saitama Prefecture. There are 60,000 people. You can get there by train from Tokyo. In Hidaka-shi there are mountains, shrines and delicious food. Hiwada mountain is Hidaka-shi’s famous mountain. It is free.
Koma shrine is Hidaka-shi’s famous shrine. There are often festivals at the shrine. You can eat a lot of delicious food.
Hidaka’s speciality/special product are chestnuts and Koma hot pot.

Although this writer was able to introduce Hidaka-shi, the information that she/he provided was very brief. For example, she/he did not provide information as to what you can do at Hiwada mountain or why it is a famous mountain.
Score (2): A good range of information is included but each piece of information may be brief and limited in detail

I came from a town called Frankston. Frankston is situated in the south of Melbourne and it takes about 50 minutes by train from Melbourne Central station to Frankston station. The ticket is 6 dollars. Since the beach in Frankston is popular, if you like swimming, you should be thinking about going to the beach in Frankston.

A town called Langwarrin is situated near Frankston. In Langwarrin there is a famous and pretty garden. In my opinion, this garden is number one place in Frankston. You need to take a bus to go to Langwarrin from Frankston. Please do not worry as the bus only costs 2 dollars. This garden is called Cruden Farm.

Since there are many activities in Frankston, I think it is a good idea to come to this wonderful town.

Compared to the writer who scored 1, this learner provided more information about the chosen destination (Frankston). For example, how long it takes to get to Frankston station, the cost of the ticket. However, he/she did not elaborate much for each piece of information.
I am from Hong Kong. Now I would like to introduce Hong Kong. Hong Kong is a small island and it takes about 3 hours from Tokyo by plane. Hong Kong is a very lively city and there are also many high-raised buildings. You can eat a lot of different cuisines in Hong Kong. There are Chinese cuisines, Korean cuisines and Japanese restaurants. The most popular cuisine is “Yum Cha” and both foreigners and locals often go to Yum Cha and eat dim sim while reading newspapers.

And, there are many shopping malls in Hong Kong and Hong Kong is very famous for shopping. The most famous shopping place in Hong Kong is Causeway Bay and it takes about 40 minutes from the airport. At Causeway Bay you can also buy clothes and electronic products. And because there are many shopping malls there are many people every day. Also, “Ocean Park” is Hong Kong’s famous theme park. There are various animals, for example, panda, shark and sea turtle. It is a very popular attraction because there are roller coasters and ferris wheel as well. I recommend everybody to go to Hong Kong.

This writer was able to provide information and elaborate on them, giving more detail. For example, she/he provided information on what kind of restaurants there are in Hong Kong and then elaborated by stating that Yum Cha is the most popular cuisine in Hong Kong.
Content: Persuasiveness

Score (0) : Gives no reasons why the place is worth visiting

私の出身ところは紹介します。それはアデライデ。そのとこは静かです。近いビルはふるいときれいです。Hahndoff は有名なまちでいろいろな Germany stores are opened. German 料理 がおいしそうですね。Glenleg Beach はたくさん seafood があります。

I will introduce the place where I come from. That is Adelaide. That place is quiet. The buildings nearby our old and beautiful. Hahndoff is a famous town with various Germany stores are opened. German food looks delicious. There are many seafood in Glenleg Beach.

This writer only stated the facts about Adelaide and did not state why Adelaide is a place worth visiting.

Score (1) Gives appropriate reasons for visiting the place but may be limited to one or two reasons

シャンハイはとても大さなおおきな町です。たくさん高いビルとおいしいたべものがあります。みなさんはいつも忙しそうです。毎日えきでちかてつで会社にいきます。シャンハイは中国のけいざいの中心だけではなく、有名なかんこうじです。一番人気を持っているのところはオリエンタルパール TV ターミナルです。一番高いビルじゃないけど、シャンハイのきれいな夜景がみられます。ちかてつ 2 をのって、かんたんにいけられます。入場券は 40 ドルです。

Shanghai is a very big city. There are many high raised buildings and delicious foods. Everybody always looks busy. Everyday using station and subway go to company. Shanghai is not only the centre of Chinese economics, it is a famous tourist destination. The most popular place is the Oriental Pearl TV Tower. It is not the highest building, but you can see Shanghai’s beautiful night view. You can get there easily by getting on the no.2 subway. The entrance fee is $40.

This writer explains that Shanghai is a famous tourist destination and introduces Oriental Pearl TV Tower explaining why it is a popular place – the night view is beautiful. The reason for going to Shanghai is very limited.
私のふるさとは雲南です。雲南省は中国の西南にあります。飛行機でも大車でも雲南に着くことができます。メルボルンの空港で飛行機に乗ります。そして、ホンコンで降ります。その後、ホンコンで昆明への飛行機に乗り換えて昆明で降ります。そうすると、私のふるさとに着けます。

雲南はとてもきれいですみやすいところです。冬はあまり寒くないし、夏は暑くないですから、観光できる場所もいっぱいあります。たとえば、万国博覧会の公園とか森とか古い城とか、いろいろな歴史的なところがあります。雲南の物価も安いので、食料と宿泊料金を含めて二万元は十分です。

皆さん、機械があったらぜひ私のふるさに旅行に来てくださいね。

My hometown is Yunnan. The Yunnan Province is situated in southwest of China. It is possible to get there by plane as well as by large car. Get on a plane at the Melbourne airport. Then, get off at Hong Kong. After that, in Hong Kong, transfer on a plane to Kunming and get off at Kunming. Then, you will be able to get to my home town.

Yunnan is a very pretty and liveable place. It is not very cold during winter and it is not hot in summer so there are many places you can go for sightseeing. There are many historic places for example, universal exposition’s parks, forests, old castle etc. Things are also very cheap in Yunnan, 20,000 yuan is enough including food accommodation expenses.

If you have the opportunity, please visit my home town.

This writer provided several reasons as to why Yunnan is a nice place to visit- climate, historical places and the price of the commodity. However, she/he did not elaborate the information further. For example, what you can do at the historical places and what are the features that attract the tourists (e.g. parks, forests).
今日はオーストラリアのメルボルンを紹介します。メルボルンはとても大きく、人口もふえている。特に、メルボルンでは西洋の文化だけではなく、アジアの文化も経験できます。メルボルンの中心の「メルボルンセンター」と言う所に行くと、色々な国の食べ物が食べられます。私が一番おすすめしあげたいレストランは「Grilled」と言うハンバーグのレストランです。あまり塩辛くなくて、ハンバーグの材料もとても新鮮で、おいしいです。その後デザートは「Tea House」と言うお茶店で、茶を飲むのをおすすめします。「Tea House」のお茶はしゅるいのが色々あるので、自分のお好みによってえらんで飲めます。そして、カスタードパンのようなあまいのもありますが、とてもおいしいです。それ以外にも、日本のレストランもありますし、中国のレストランもありますし、色々ありますし、色々あります。もし、パスタとかピザなどのイタリアの食べ物が食べたがっていたら、「Lygon St」に行ってみて下さい。本当のイタリア人が料理しますので、食べたらまるでイタリアの旅行をしている気分になるかもしれませんが。それにジェラートのお店もあります。

オーストラリアナンバーワンの大学「メルボルン大学」は、まるで「ハリーポッター」に出るホグワーツのようにすてきで、大きすぎています。だから、まいごになるかもしれませんから、気をつけて下さい。大学に行く前にネットで地図を見たら問題ないと思います。

メルボルンではきれいでもかわいい物がたくさんありますが、値段を見たら、すっかりびっくりするかもしれません。メルボルンは日本より物価が高すぎています。だから安い値段できれいなものがほしけていたら、「DFO」のようなアウトレットに行けばいいんです。「DFO」はほかのお店よりずっと安いです。

こんなにおいしい食べ物がたくさんあって、あそぶ事もたくさんあるメルボルンに行ってみてください！

Today I will introduce Melbourne, Australia. Melbourne is very big and the population is growing. Especially in Melbourne not only you can experience the western culture but you can also experience the Asian culture. If you go to a place called “Melbourne Centre” in the centre of Melbourne, you can eat food from various countries. The restaurant that I would like to recommend the most is a hamburger restaurant called “Grilled”. It is not that salty and the hamburger ingredients are very fresh and it is delicious. After that I recommend for you to drink tea as a dessert at a tea shop called “Tea House”. There is a lot of variety of teas at the “Tea House” so you can choose depending upon your taste. And
there is also a sweet things like custard bun but it is very delicious. Other than that, there are Japanese restaurants, Chinese restaurants and other various ones. If you feel like eating Italian food like pasta or pizza, you should go to “Lygon St”. Cooking is done by the native Italian, so when you eat, you may feel like as if you are travelling in Italy. Moreover, there is also a Gelato shop.

Australia’s number one university, “Melbourne university” is beautiful like Hogwarts in Harry Potter and is too large. Therefore, you may get lost so be careful. If you look at the map on the internet before you go to the university you should not have a problem.

There are a lot of beautiful and cute things in Melbourne but is you look at the price, you may get completely surprised. The prices in Melbourne are too high compared to Japan. Therefore, if you want a good that are cheaper, you should go to outlet like “DFO”. DFO is a lot cheaper than other stores.

Please pay a visit to Melbourne where there are such delicious foods and various places to visit!

For each information provided, this writer gave reasons as to why the recommended place is worth visiting. For example, he/she provided extra information that Italian food on Lygon Street is authentic which would convince the reader to try and go to that restaurant.
Coding examples for each rating scale.

Please note: The examples were taken directly from the learners’ writing script. Therefore, some sentences are not necessarily written correctly. The English translation was translated as close as possible but some errors were unable to be translated.

Structure: Coherence

Score (0): Entire text difficult to follow

There is no connection from one sentence to another in this piece of writing. The entire text is difficult to follow.
Welcome! Melbourne is Australia’s number one cultural centre. There is Eureka Tower in the city of Melbourne. Eureka Tower is very high and from the tower you can see the Melbourne CBD. Other than the Eureka Tower, we eat various cultural food and watch theatre at the Melbourne Arts Centre. And, Melbourne’s Luna Park is very enjoyable. We eat popcorn at Luna Park and go on the ride. Melbourne is beautiful and I think it is a very interesting place. You can come to Melbourne by train. It sounds very convenient.

This writer attempted to introduce the main attraction in Melbourne. However, some sentences do not flow well from one another. For example, from Eureka Tower, he/she jumps straight into Luna Park without any connecting sentence such as “if you feel like going on a ride after a scenic view at Eureka Tower, Luna Park is the place to go”.

Score (1) : Several ideas are not located where they should be, presents information with limited organisational flow; several ideas not linked smoothly to each other
Melbourne is a city that has a long history in Australia. In the past, since Australia’s government could not decide whether to make Sydney or Melbourne the capital city, so they created a new town and it seems that Canberra was formed. Canberra is still Australia’s capital city.

Melbourne is very lively city. There are many overseas students because there are many famous universities such as Melbourne University, RMIT and Monash University in the city. Therefore, you can hear different languages and slangs while walking the city.

The lifestyle in Melbourne seems to be relaxing compared to other countries, therefore there’s more time every day. There is a coffee culture and brunch culture among the young people here. Other than that, there are many old buildings and parks in the city.
There are trams and trains in the city of Melbourne and you do not need to drive a car. It is difficult to park in the city and it is very expensive. However car is essential if you want to go to the outer suburbs. The beaches around Melbourne are very beautiful and during summer there are also many people.

However, everything in Melbourne is expensive. Compared to other countries, it seems like it would be difficult if you don’t have a lot of money.

Coherence was seen within each paragraph in this piece of writing, however, sometimes there was no link between one paragraph and another.
Melbourne is Australia’s city. It is a very big and interesting city and there are many activities that you can do. In order to go to Melbourne you can use train, tram or plane. When you do sightseeing in Melbourne, it seems that many people use train and trams. This is because Melbourne has the biggest tram network.

Many tourists come to Melbourne because there are many things you can do. I believe the most popular activity is the Eureka Tower. Eureka Tower has 88 storey and it is the highest building in the Southern Hemisphere. It is on the top floor and has a beautiful and magnificent view. You can have the view of the whole Melbourne city. I think the night view is better.

If you like the splendid view, you should go to Melbourne Star. The view is beautiful. Melbourne Star is a very big Ferris wheel. Melbourne Star and Eureka Tower costs about $20.
If you don’t have a lot of money, you should go to a café in Melbourne. Melbourne’s café are very famous because it is very cool and interesting. Finally, there are also many museums in Melbourne. For example, museums and Melbourne zoo. Usually there cost about $15 and museums are free.

This writer was able to connect the information from one sentence to another as well as from one paragraph to another. For example, after introducing Eureka Tower in the second paragraph, the writer links to the next paragraph by starting off with the statement “if you like the splendid view” and then introduces another attraction in Melbourne (i.e. Melbourne Star).
If you like history, Melbourne museum is a nice place. Since the museum is situated in Carlton, it would be convenient to go there by bus. The price of the museum ticket is not that expensive. Sometimes there is a special exhibition at the museum. If that is the case, I recommend you to have a look. After you have seen everything at the museum, you can go to the park next to the museum called “Carlton Gardens”. You can see beautiful flowers and trees all year round. Once a year there is an Interflora Flower Show at Carlton Gardens.

There is no paragraph in this piece of writing. Ideas are not split.
Beautiful Melbourne!

Melbourne is small but a very beautiful city. There is a Eureka Tower in Melbourne. Eureka Tower is a high building in Australia. From the tower, you are able to see the city.

There is a Moomba festival around March every year. People listen to the music in the city and watch water sports at Yarra River. The city is very lively. Travelling is easy because there is a tram, train and plane in Melbourne.

These sentences need to be merged into one paragraph as it is explaining about the same topic, Melbourne Zoo.
There are clearly distinguished paragraphs and related ideas are appropriately grouped into paragraphs.

St. Kilda

St. Kilda is located in the south of Melbourne. It takes about 15 minutes from Flinders Station to St. Kilda. You will be able to get there if you take the No. 96 tram.

There is a very famous beach in St. Kilda. If you like to swim, it is a good idea to go to this beach. Other than the beach, a fun place that you can go is Luna Park. Luna Park is an old amusement park. Luna Park has many tourist attractions, there is also a roller coaster called a mad mouse. The entrance fee is 25 dollars.

There is also a historical site in St. Kilda. For example, a theatre called National Arc is very popular. This year, the comedy festival’s performance was shown.

Paragraphs are clearly distinguished and ideas are grouped appropriately.
Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Kikuchi, Yuka

Title:
Writing development of second language learners of Japanese at an Australian university

Date:
2017

Persistent Link:
http://hdl.handle.net/11343/192974

File Description:
Writing Development of Second Language Learners of Japanese at an Australian University

Terms and Conditions:
Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.