

THE COHESION FACTOR: A STUDY OF JAPANESE JUNIOR HIGH SCHOOL WRITING

by

DENEYS LAURENCE COOMBE

**Submitted in accordance with the requirements
for the degree of**

**MASTER OF ARTS WITH SPECIALIZATION IN TESOL
(TEACHING ENGLISH TO SPEAKERS OF OTHER LANGUAGES)**

at the

UNIVERSITY OF SOUTH AFRICA

Student Number: 7497334

SUPERVISOR: DR R A SCHEEPERS

JUNE 2015

Notes for the examiners

Although Unisa is a South African institution where English (as opposed to American) spelling is the norm, this dissertation was written in Japan, which follows American spelling rules. Almost all the computers used employ American spelling as a default setting. As a result, American spelling rules will be followed. The writer appreciates your understanding on this matter, and hopes that in this case, pragmatism will triumph over principle. Particular words include: *color, endeavor, practice* (as a verb), *dialog, and program*

The first chapter provides some general information on ELT in Japanese schools. While this information is not directly related to the main focus of this study (namely cohesion), it was nevertheless deemed necessary to offer readers outside Japan some background information on English teaching and learning in Japan.

Abstract

This study compared cohesive devices in texts written by Japanese second-year junior high school learners with those in texts that appeared in the textbook they were studying. The purpose of the study was to determine which cohesive devices were being used in the textbook and which were used in the learners' writing. The study used both quantitative and qualitative analysis. The quantitative analysis began by determining whether there was any significant difference between the textbook readings and the learners' writing in terms of the frequency of cohesive devices. It then examined the kinds of devices that were used by both groups of texts. The qualitative analysis compared the patterns of reiteration in two textbook readings with those in a sample of six student texts of different levels of success. The results showed no significant differences between the student texts and the textbooks in terms of the overall frequency of cohesive devices. Among the individual devices, however, there was a significantly higher frequency of ellipsis and synonyms in the textbook readings than in the student texts. There was also a significantly higher frequency of conjunction and reference in the student texts relative to the textbook readings. In all other devices, there was no significant difference between the textbook readings and the student texts. The qualitative study revealed the importance of strong opening sentences, reinforcement of the main topic through repetition, as well as of linking new topics with the main topic in the textbook readings. However, the presence of these features varied in the selected student texts. Accordingly, stronger texts contained all these features, average texts contained some of them, and weaker texts contained few or none. This study consequently supports other studies that have shown that the way in which cohesive devices are used is far more important in determining text quality than the number of devices used. The findings of this study showed the strengths and weaknesses in the students' writing, and highlighted the need for a greater awareness of cohesion by focusing more on sentence building, and the use of a greater variety of cohesive devices.

Key words: cohesion; high school; writing; Japanese; cohesive devices; elementary level;
EFL

Acknowledgements

On the academic side:

First and foremost, to my supervisor, Dr Ruth Scheepers: for her thorough feedback, patience, and assistance; for going beyond the call of duty on many occasions, for on more than one occasion saving me from giving up altogether, and for staying around long enough to see this dissertation through to its completion.

To Professor Brenda Spencer: for also rescuing me from the brink, and for her help in many areas both academic and administrative.

To all members, both academic and administrative of Unisa English Faculty: this dissertation would not even have been possible had it not been for the amazing support and encouragement of TESOL team during my Honors years. I am particularly grateful to Dr Rita Ribbins and Dr Christine Marshall who, during those years, were accommodating and understanding, and always available to answer my many queries.

To my colleague, Arnold Tiempo, for his much valued feedback and encouragement.

To all colleagues, past and present, for providing me with books and other reference material, for their patience in listening to my woes, and for giving me time to work. Their professionalism and dedication have always been inspirational.

To Alexa Barnby for her final editing and to Gerry Barnby for his work in formatting this document; your skill, advice and contributions were invaluable.

To Chris van der Leij, for his advice and encouragement at an important juncture.

On the non-academic side:

Most importantly, to my dear long-suffering wife, Hitomi Kimura, who shouldered much of the domestic burden while I worked on my days off and some evenings, and who did her best to give me the time to work. Thank you so much for everything.

To my two sons, Thomas (9) and Ewart (6), who went without Papa on many days off and other holidays. We have much catching up to do.

To my parents, Ed and Barbara Coombe, for their constant interest and encouragement over the years.

To the management and staff of 'SB' coffee shop, Shonandai, whose kindness and friendliness helped to ease much of the loneliness of writing this dissertation.

Table of contents

Declaration of own work.....	i
Notes for the examiners.....	ii
Abstract.....	iii
Acknowledgements.....	v
Table of contents.....	vii
List of diagrams.....	xiii
List of tables.....	xiv
Chapter 1: Introduction.....	1
1.1 INTRODUCTION.....	1
1.2 BACKGROUND TO THE STUDY.....	1
1.2.1 English teaching in Japan.....	2
1.2.2 Defining successful or good quality writing.....	4
1.3 WHAT IS COHESION?.....	5
1.3.1 Types of cohesion.....	5
1.3.1.1 Grammatical cohesion.....	6
1.3.1.2 Lexical cohesion.....	6
1.3.2 Composition structure: differing approaches.....	8
1.3.3 Same skills, different structure.....	10
1.3.4 Attention to audience.....	11
1.4 DEFINING THE RESEARCH PROBLEM.....	13
1.5 RESEARCH AIMS.....	14
1.6 ASSUMPTIONS AND RATIONALE.....	14
1.7 THE RESEARCH QUESTIONS.....	16
1.8 THESIS STATEMENT AND DELINEATIONS.....	17
1.8.1 Delineations.....	17
1.8.2 Definition of terms.....	18
1.9 SUMMARY.....	18
Chapter 2: Review of the Literature.....	21
2.1 INTRODUCTION.....	21
2.2 COHESION AND COHERENCE.....	24
2.3 THE KNOWN-NEW CONTRACT.....	31
2.4 STUDIES IN COHESION.....	32

2.4.1	Cohesion and language proficiency	33
2.4.2	Types of cohesive devices in EFL writing	37
2.4.2.1	Demonstratives, conjunctions and sentence transition	39
2.4.2.2	Repetition.....	40
2.4.2.3	Synonyms	44
2.4.3	Cohesion errors and first language (L1) transfer.....	46
2.4.4	Exophoric referencing.....	51
2.4.5	Focus on cohesion in Japanese high schools	52
2.5	SUMMARY AND CONCLUSION	55
Chapter 3: Method.....		59
3.1	INTRODUCTION	59
3.2	RESEARCH AIMS AND QUESTIONS	59
3.3	RESEARCH DESIGN	60
3.4	FOCUS OF THE STUDY.....	60
3.5	CONTEXT OF THE STUDY	60
3.6	PARTICIPANTS	63
3.7	SAMPLING METHOD.....	64
3.8	DATA.....	64
3.8.1	The student texts (STs).....	64
3.8.2	The textbook readings (TBRs).....	65
3.9	PREPARATION OF DATA	67
3.10	FRAMEWORKS FOR THE STUDY.....	69
3.10.1	Halliday and Hasan's taxonomy of cohesive devices (1976)	70
3.10.1.1	Grammatical cohesive devices	70
3.10.1.2	Lexical devices	75
3.10.2	Adaptation and simplification of Halliday and Hasan's (1976) taxonomy.....	78
3.10.3	Coding of the data	80
3.10.4	Hoey's (1991) patterns of lexis model	81
3.10.4.1	Simple repetition	83
3.10.4.2	Complex lexical repetition.....	83
3.10.4.3	Paraphrase	83
3.10.4.4	Superordinate, hyponymic, and co-reference repetition.....	84
3.10.4.5	Other types of repetition	84
3.11	DATA ANALYSIS.....	85

3.11.1	Quantitative data analysis.....	86
3.11.2	Qualitative data analysis.....	87
3.12	ETHICAL CONSIDERATIONS	90
3.13	SUMMARY AND CONCLUSION	91
Chapter 4: Results and Discussion of the Findings.....		93
4.1	INTRODUCTION	93
4.1.1	Reiteration of main research questions	93
4.2	FINDINGS RELATING TO QUESTION 1	93
4.3	FINDINGS RELATING TO RESEARCH SUB QUESTION 2	96
4.3.1	Grammatical devices	97
4.3.1.1	Conjunctions	98
4.3.1.2	Reference	99
4.3.2	Lexical cohesive devices	103
4.3.2.1	Repetition.....	103
4.3.2.2	Synonyms	105
4.3.2.3	Superordinates	106
4.3.2.4	Collocation	108
4.3.3	Summary of quantitative results	108
4.4	FINDINGS RELATING TO QUESTION 3	109
4.4.1	A variety of cohesive devices in cohesive chains	112
4.4.2	The opening sentence and repetition links	113
4.4.3	The main topic or theme	116
4.4.4	Clusters of repetition.....	117
4.4.5	Sentences of varying length	121
4.4.6	A final sentence that links back to the middle and beginning	123
4.5	FINDINGS RELATING TO QUESTION 4	124
4.5.1	Student texts: a note on grading.....	124
4.5.2	Criteria for analysis	125
4.5.3	Results of the analysis.....	125
4.5.4	ST 1 Tom Sawyer.....	126
4.5.4.1	A variety of cohesive devices used in cohesive chains	127
4.5.4.2	The opening sentence and repetition links	129
4.5.4.3	The central topic or theme	130

4.5.4.4	Clusters of repetition.....	130
4.5.4.5	Sentences of varied length	131
4.5.4.6	The final sentence	132
4.5.4.7	Summary	133
4.5.5	ST 2 The Jumble Sale	133
4.5.5.1	A variety of cohesive devices used in cohesive chains	134
4.5.5.2	The opening sentence and repetition links	136
4.5.5.3	The central topic or theme	137
4.5.5.4	Clusters of repetition.....	137
4.5.5.5	Sentences of varied length	139
4.5.5.6	The final sentence	141
4.5.5.7	Summary	142
4.5.6	ST 3 Escape	143
4.5.6.1	A variety of cohesive devices used in cohesive chains	143
4.5.6.2	The opening sentence and repetition links	145
4.5.6.3	The central topic or theme	146
4.5.6.4	Clusters of repetition.....	146
4.5.6.5	Sentences of varied length	147
4.5.6.6	The final sentences.....	148
4.5.6.7	Summary	148
4.5.7	ST 4 Phantom of the Opera.....	149
4.5.7.1	A variety of cohesive devices used in cohesive chains	149
4.5.7.2	The opening sentence and repetition links	150
4.5.7.3	The central topic or theme	151
4.5.7.4	Clusters of repetition.....	151
4.5.7.5	Sentences of varied length	152
4.5.7.6	The final sentence	153
4.5.7.7	Summary	153
4.5.8	ST 5 The Scarf.....	154
4.5.8.1	A variety of cohesive devices used in cohesive chains	154
4.5.8.2	The opening sentence and repetition links	155
4.5.8.3	The central topic or theme	156
4.5.8.4	Clusters of repetition.....	156
4.5.8.5	Sentences of varied length	157
4.5.8.6	The final sentence	158
4.5.8.7	Summary	158

4.5.9	ST 6 Sangokushi	159
4.5.9.1	A variety of cohesive devices used in cohesive chains	159
4.5.9.2	The opening sentence and repetition links	160
4.5.9.3	The central topic or theme	161
4.5.9.4	Clusters of repetition	161
4.5.9.5	Sentences of varied length	162
4.5.9.6	The final sentence	163
4.5.9.7	Summary	164
4.5.9.8	Summary of qualitative analysis	165
4.6	Findings related to Research Questions 5 and Questions 6:	165
4.6.1	Exophoric reference.....	168
4.6.2	Vague or ambiguous referencing.....	170
4.6.3	Conjunction	172
4.6.3.1	Model sentences from the textbook.....	172
4.6.3.2	Limited active knowledge of conjunctions	174
4.6.4	Redundant repetition	175
4.7	SUMMARY AND CONCLUSION	176
Chapter 5: Conclusion.....		179
5.1	INTRODUCTION	179
5.2	SUMMARY AND DISCUSSION OF THE FINDINGS	179
5.2.1	Findings for Research Question 1	179
5.2.2	Findings for Research Question 2	180
5.2.2.1	Ellipsis.....	180
5.2.2.2	Synonyms	181
5.2.2.3	Conjunction	181
5.2.2.4	Reference	182
5.2.2.5	Repetition and superordinates.....	182
5.2.3	Findings for Research Question 3	185
5.2.4	Findings for Research Question 4	186
5.2.5	Findings for Research Question 5	187
5.2.6	Findings for Research Question 6	187
5.3	CONTRIBUTIONS OF THIS STUDY	188
5.4	LIMITATIONS.....	190
5.5	RECOMMENDATIONS AND SUGGESTIONS FOR PRACTICAL APPLICATION	191

5.6	SUGGESTIONS FOR FURTHER RESEARCH	196
5.7	CONCLUDING COMMENTS	197
	BIBLIOGRAPHY	199
	Appendix 1: Summary of cohesion and coding scheme in Halliday and Hasan (1976:333–338)	204
	Appendix 2: Writing assignment.....	211
	Appendix 3: Letter of consent	213
	Appendix 4: Sample of student text coding.....	214
	Appendix 5: Reading texts from Progress in English Book 2 (Flynn 2010a)	215
	Appendix 6: Raw scores and complete tabulations	226
	Appendix 7: Superordinates found in the TBRs	230
	Appendix 8: An example of exercise on improving writing.....	232
	Appendix 9: Letter of ethical clearance	234

List of diagrams

Diagram 3.1:	Sample grid of reiteration patterns	88
Diagram 3.2:	The links between repeated words in Text 10 <i>The Lost Key</i>	89
Diagram 4.1:	Patterns of reiteration for TBR 1 <i>A Beautiful Symbol of Friendship</i>	114
Diagram 4.2:	Patterns of reiteration for TBR 2 <i>Abraham Lincoln</i>	115
Diagram 4.3:	Bonding sentences in TBR 1 <i>A Beautiful Symbol of Friendship</i>	121
Diagram 4.4:	Patterns of reiteration for ST 1 <i>Tom Sawyer</i>	129
Diagram 4.5:	Bonded sentences 4, 9 and 10 in ST 1 <i>Tom Sawyer</i>	131
Diagram 4.6:	Patterns of reiteration for <i>The Jumble Sale</i>	137
Diagram 4.7:	Bonded sentences 1, 4 and 9 in ST 2 <i>The Jumble Sale</i>	140
Diagram 4.8:	Bonded sentences 1 and 8 in <i>The Jumble Sale</i>	140
Diagram 4.9:	Patterns of reiteration for <i>Escape</i>	145
Diagram 4.10:	Bonded sentences 1 and 4 in <i>Escape</i>	147
Diagram 4.11:	Patterns of reiteration for <i>Phantom of the Opera</i>	150
Diagram 4.12:	Linked sentences 3 and 6 in <i>Phantom of the Opera</i>	152
Diagram 4.13:	Patterns of reiteration for <i>The Scarf</i>	155
Diagram 4.14:	Bonded sentences 3 and 7 in <i>The Scarf</i>	158
Diagram 4.15:	Patterns of reiteration for <i>Sangokushi</i>	161
Diagram 4.16:	Sentence links 4, 5 and 6 in <i>Sangokushi</i>	163

List of tables

Table 3.1:	Grammar structures covered in the first semester in J2 when the essays were written	62
Table 3.2:	Summary of cohesive types identified in this study.....	79
Table 3.3:	Sample of classification of cohesive types.....	80
Table 4.1:	Number of cohesive devices identified in each corpus	94
Table 4.2:	Log-likelihood results for all devices.....	96
Table 4.3:	Log-likelihood results for grammatical devices	97
Table 4.4:	Words and phrases referred to by the demonstratives in the STs	100
Table 4.5:	Words and phrases referred to by demonstratives in the TBRs	101
Table 4.6:	Definite article referents that occurred two or more times.....	102
Table 4.7:	Log-likelihood results for lexical devices	103
Table 4.8:	Repeated words according to parts of speech log-likelihood table.....	104
Table 4.9:	Type-to-token ratio for superordinates used in the STs and TBRs	106
Table 4.10:	Superordinates used more than once in TBRs and STs	107
Table 4.11:	Lexical chains and cohesive types in TBR B <i>A Beautiful Symbol Of friendship</i>	112
Table 4.12:	Lexical chains and cohesive types in TBR 2 <i>Abraham Lincoln</i>	113
Table 4.13:	Repetition clusters in TBR 1 <i>A Beautiful Symbol of Friendship</i>	118
Table 4.14:	Repetition clusters (2) in TBR 2 <i>Abraham Lincoln</i>	119
Table 4.15:	Repetition clusters (1) in TBR 2 <i>Abraham Lincoln</i>	120
Table 4.16:	Rater scores for the six selected student texts	125
Table 4.17:	Lexical chains and cohesive types (1) in ST 3 <i>Tom Sawyer</i>	127
Table 4.18:	Lexical chains and cohesive types (2) in ST 3 <i>Tom Sawyer</i>	128
Table 4.19:	Repetition clusters in ST 1 <i>Tom Sawyer</i>	130
Table 4.20:	Lexical links to the penultimate sentence in ST 1 <i>Tom Sawyer</i>	132
Table 4.21:	Lexical chains and cohesive types (1) in ST 2 <i>The jumble sale</i>	134
Table 4.22:	Lexical chains and cohesive types (2) in ST 2 <i>The Jumble Sale</i>	135
Table 4.23:	Lexical chains and cohesive types (3) in ST 3 <i>The Jumble Sale</i>	135
Table 4.24:	Repetition clusters (1) in ST 2 <i>The Jumble Sale</i>	137

Table 4:25:	Repetition clusters (2) in ST2 <i>The Jumble Sale</i>	138
Table 4:26:	Links to the penultimate and ultimate sentences in ST 2 <i>The Jumble Sale</i>	141
Table 4:27:	Lexical chains and cohesive types (1) in ST 3 <i>Escape</i>	143
Table 4:28:	Lexical chains and cohesive types (2) in ST 3 <i>Escape</i>	144
Table 4:29:	Lexical chains and cohesive types (3) in ST 3 <i>Escape</i>	144
Table 4:30:	Lexical chains and cohesive types (1) in <i>Phantom of the Opera</i>	149
Table 4:31:	Lexical chains and cohesive types (2) in <i>Phantom of the Opera</i>	149
Table 4:32:	Lexical chains and cohesive types (3) in <i>Phantom of the Opera</i>	150
Table 4:33:	Repetition clusters in <i>Phantom of the Opera</i>	151
Table 4:34:	Lexical chains and cohesive types (1) in <i>The Scarf</i>	154
Table 4:35:	Lexical chains and cohesive types (2) in <i>The Scarf</i>	154
Table 4:36:	Near-repetition cluster in <i>The Scarf</i>	156
Table 4:37:	Lexical chains and cohesive types for <i>Sangokushi</i>	160
Table 4:38:	Near-repetition clustering in <i>Sangokushi</i>	162
Table 4:39:	A summary of how the six STs compare to the TBRs	165
Table 4.40:	Total number of cohesive errors in the STs (excluding repetition) in order of the number of errors	166

Chapter 1: Introduction

1.1 INTRODUCTION

This study compared the cohesive devices found in a textbook used in a Japanese junior high school with those found in the writing of learners who were using this book at the time. Using the textbook readings as a model, the study aimed to determine the extent to which the learners' writing differed from or was similar to the reading sections used in the textbook. The study was based on the premise that the types and frequency of cohesive devices found in the textbook used by the learners, as well as the way in which these devices were used, provided a model and a yardstick against which the writings of the learners using the textbook could be measured and compared. Having measured and accounted for the use of cohesive devices in the textbooks, and compared this to the learners' writing, it would be possible to gain a better understanding of how the use of cohesion affected the quality of the learners' writing.

This chapter provides the background to the study. It explains the situation in Japan regarding the teaching of English composition in Japanese junior and senior high schools. After briefly examining some of the challenges and benefits of teaching composition, it discusses the similarities and differences in the structure of English and Japanese texts. This leads to a brief introduction to cohesion in English, and its importance in the teaching and assessment of English compositions in a Japanese school context. The stage is thus set for defining the research problem and the research questions that arise.

1.2 BACKGROUND TO THE STUDY

As a South African with around eight years' teaching experience in Japan, I was fortunate to find a job at a private junior high school near Tokyo in 2006. In the course of this work, I became increasingly aware of the power of composition as a means of allowing learners to

express themselves. From this arose an interest in how writing is taught and learnt in Japanese English classes, particularly at the school where the research took place.

After some discussion with teachers, it was found that, while composition was becoming an imperative in high school, many teachers had certain reservations about teaching it. For some, there was the simple problem of having to grade 180 essays, given their already busy schedules, while for others there was the question of how best to assess essays. Should they opt for a holistic approach, namely, giving one grade for the essay as a whole, or should they award separate grades for individual criteria? Should they focus on grammatical details or the structure of the essay as a whole? What is it exactly that makes one text better than another? With these questions in mind, I initially tried to cast my net as wide as possible. However, deeper probing, and an increased awareness of what teachers were doing in class, led to the realization that a study which tried to cover every aspect of writing comprehensively would be neither necessary nor feasible. Many teachers were already modeling texts and attending to the structure of different genres. Some teachers were also following particular writing processes. Other teachers were acutely aware of grammar mistakes. Cohesion, however, seemed to fall somewhere between the extremes of text structure on the one hand, and sentence accuracy on the other. While there was a definite awareness of many cohesive devices, there still appeared to be the need for a better understanding of which devices were being used by students at different proficiency levels. The end result was a study that focused on cohesion.

1.2.1 English teaching in Japan

The major goal of many learners in Japan is to enter a prestigious university. This achievement can have a major impact on their employment prospects. In a country where, until a few years ago, lifetime employment (that is, employment in the same company for a person's entire working life) was the norm, good results at high school have always had

enormous implications. Within this context, the university entrance English test can be a major factor in determining which university a student may attend. In certain situations, where there is little or no difference between mathematics and science scores, English proficiency may in fact be the deciding factor in determining admission to a particular university.

Until about 20 years ago, English was taught in Japanese high schools almost exclusively by the grammar-translation method. This approach entails the translation of sentences from one language to another, with very little focus on communicative ability. Even though guidelines for the teaching of communicative English were introduced by the Ministry of Education almost 20 years ago in 1994 (Takagi 2001), the bulk of the university entrance tests remain grammar-translation based (Takagi 2001). Furthermore, most grammatical structures are taught as discrete units, often resulting in a deep and detailed understanding of English grammar, but little communicative or discourse competence. As far as writing is concerned, while learners are able to recognize grammatical structures at sentence level, they are often confused by longer texts, and find it difficult to synthesize sentences into paragraph form (Cunningham [sa]). This focus on grammar translation continues, even as teachers branch out into other approaches and methods of teaching. To give an example, at our school, students are taught English in two different classes, namely, English 1 and English 2. The former is almost always taught by a Japanese teacher, and is concerned mainly with introducing new grammar structures. A large part of this is done through the medium of Japanese, and a substantial percentage of the questions in the tests require some kind of translation, mainly from Japanese to English. The English 2 component is focused more on applying the learned grammar to reading, writing and speaking. Nevertheless, one still sees learners tackling a text by translating it sentence by sentence.

In recent years, many universities have begun to include a free-writing component in their entrance tests, meaning that both learners and teachers have had to learn the art of paragraph writing. This has become an added burden to both teachers and learners. Writing is still regarded by many as sentence-level work. This is reflected in Gates's (2003) study in which most of the exercises in the textbooks he looked at did not allow for much free writing. Furthermore, the introduction of more communicative methods has seen a focus on speaking and listening, at the expense of extended reading and writing.

1.2.2 Defining successful or good quality writing

Many native English-speaking teachers may themselves be hard-pressed to explain precisely to a learner why his or her composition has been awarded a lower grade than that of his peers. Remarks such as 'it was more interesting,' or 'your style needs work' are often used. But defining precisely what constitutes an interesting essay or good style can be very difficult. Indeed, I, a native speaker of English, struggled to put my finger on some of these points until relatively recently. Some of the main points would include the following:

- **The structure of the essay.** This includes introductions and conclusion, how the main topics, themes or narrative are arranged.
- **Content.** This refers to the kind of information that the writer provides to the reader. A good text usually provides just the right amount of information: not so little that the reader has to infer what is being said, and not so much that the reader is left bored or overwhelmed.
- **The flow of the language.** In a successful or high quality text, the sentences seem to flow naturally from one to another. On the other hand, in a text of poor quality, the grammar may be perfect, but the sentences seem jarring to the reader. Precisely why one text may seem more natural than another is difficult to define, and is a question that this study aims to address.

- **Accuracy.** This refers to accuracy in both grammar and spelling. There appears to be no discernible link between the ‘top down’ or holistic approach of assessing text structure and attention to audience on the one hand (the text's fluency), and the 'bottom up' focus on grammar and spelling on the other (namely, its accuracy).

What is needed is a more effective way of explaining to learners why certain sentences do not sound natural, or are too wordy, even if the sentence is grammatically correct. They need to be shown how to eliminate unnecessary words and to know how to reiterate ideas without necessarily repeating exactly the same words. Furthermore, they need to understand when repetition may or may not be appropriate. All of these problems can be explained, to some extent at least, in terms of cohesion.

1.3 WHAT IS COHESION?

According to Richards and Schmidt (2002), cohesion refers to “the grammatical and/or lexical relationships between the different elements of a text. This may be between sentences or between different parts of a sentence” (2002:86).

This can be achieved through the use of a number of different cohesive devices, some of which are enumerated in the next section (for an exhaustive list, please refer to Appendix 1). Some of the questions that will be explored are concerned with the extent to which the quality of the text is reflected in the frequency of cohesive devices, and in the way cohesive devices are used in the text.

1.3.1 Types of cohesion

Halliday and Hasan (1976) divide cohesion into two broad types, namely, grammatical cohesion and lexical cohesion. Within these two broad types, the devices are classified as follows:

1.3.1.1 Grammatical cohesion

Grammatical cohesion includes the following:

- **Reference:** personal pronouns (*I, me, you*), demonstratives (*this, that, those, these*), comparative signals (*same, different, other, better, more*).
- **Conjunction:** A. additive (*and, or, besides, by contrast, furthermore, likewise, on the other hand, etc.*); B. adversative (*yet, but, though, however, instead, etc.*); C. causal (*so, then, for, because, to this end, etc.*); D temporal (*then, next, first, later, finally, etc.*).
- **Ellipsis:** parts of the sentence that are left out or 'understood' (Tim ate two pancakes and Shelley (*ate*) four; I am older than you (*are*). He asked me to go with him, but I didn't (*go with him*).
- **Substitution:** words that are substituted for other structures (Tim served pancakes, but I didn't want **any**; Pat thinks our test will be easy, but I don't think **so**. In the first example, *any* is substituted for *pancakes*, while in the second, *so* takes the place of *our test will be easy*.

1.3.1.2 Lexical cohesion

Lexical cohesion refers to the way in which words and their meanings combine to give cohesion to a text. There are three main groups:

1. **Reiteration** (repetition of the same word; synonyms and near synonyms; superordinate words [words that define a category, for example animal is a superordinate of dog]; general words [these are similar to superordinates but encompass far more general categories, for example the word thing may refer to a number of categories of objects, situations or ideas]; other examples are stuff, objects, issue, etc. (My own examples).

2. **Collocation** (words that generally co-occur) *boy/girl; stand up/sit down; order/obey; king/crown, etc.*)
3. **Parallelism**. This device 'suggests a connection simply because the form of one sentence or clause repeats the form of another' (Cook 2004:15). For example, Winston Churchill's 'We'll fight them on the beaches' speech, or more recently, Barack Obama's 'Yes we can' speech repeat the same form numerous times (my own examples). According to Cook, this device is often used in 'speeches, prayers, poetry and advertisements, prayers and in poetry. It can have a very strong emotional effect on the reader, and is also a useful aide-memoire' (2004:15).

According to Cook (2004:15), parallelism does not necessarily have to involve grammatical items. There may be sound parallelisms, namely, repeating rhymes and rhythms. There is also semantic parallelism, in which two sentences are linked by their similar meaning. This is often used for comedic effect:

E's not pinin'! 'E's passed on! This parrot is no more! He has ceased to be! 'E's expired and gone to meet 'is maker! 'E's a stiff! Bereft of life, 'e rests in peace! If you hadn't nailed 'im to the perch 'e'd be pushing up the daisies! ... THIS IS AN EX-PARROT!!

In this classic piece of comedy from Monty Python (www.montypython.net), the comedic effect lies in the speaker echoing his initial utterance, namely, that the parrot is dead, but using a number of alternative expressions.

Halliday and Hasan's (1976) classification of cohesive devices is often seen as a definitive work in the study of cohesion. As such, it has been used in a number of studies (Hinkel 2001; Olateju 2006; Ramasawmy 2004). However, while most studies have focused on grammatical cohesion, lexical cohesion has not been covered in quite as much depth

(McGee 2009). Following on from the work of Halliday and Hasan (1976), and later Hasan (1984), scholars such as Hoey (1991) and Winter (1994) expanded the study of lexis in writing. Indeed, Hoey (1991) makes a very strong argument for the case that most cohesive elements in a text, with the exception of reference, are in fact lexical. It has also been noted that in texts written by learners of a language, most cohesive devices are grammatical, while more advanced texts tend to make more use of lexical cohesion (McGee 2009). The present study set out to find out the extent to which these previous findings were true.

1.3.2 Composition structure: differing approaches

One of the most important questions concerning the teaching of writing in a foreign language concerns the extent to which writing skills in the first language can enhance or hinder the development of skills in the target language. In many respects, the move towards more communicative activities in English language teaching (ELT) in Japan can be seen as enhancing what Cummins (2008) refers to as basic interpersonal communication skills or BICS, whereas the teaching of writing requires the development of what is referred to as cognitive/academic language proficiency, or CALP (Cummins 2008). When learners write compositions in their native language, they are in fact exercising CALP, or cognitive skills, rather than their BICS, or their interpersonal communication skills, meaning that they are applying not simply language skills, but also their knowledge and understanding of the writing process as well the topic about which they are writing. As Kimball (1996:59) points out, one of the main obstacles to academic writing in Japanese schools and universities is not so much a lack of language skill as a lack of 'competence in writing strategy'. The focus on the overall structure of a text involves the use of CALP in order to recognize the discourse convention of the text. While learners are used to operating on a communicative level, they now need to start applying their cognitive skills, some of which may be transferred from their native tongue. Even within writing itself, there may well be a

BICS/CALP distinction at play, since the kind of writing done in junior high schools may focus on more personal topics, while in senior high school and universities, academic writing will require more CALP.

Japanese schools are divided into elementary, junior high and high schools. Elementary school lasts from the South African equivalent of Grades 1 to 6. Junior high school lasts three years from the South African Grades 7 to 9, while senior high school lasts from Grades 10 to 12. In this study, the junior high grades will be referred to as Junior 1 to 3, and the senior high grades will be referred to as Senior 1 to 3. In most cases, these phases are catered for by different schools. Often learners in Junior 3 need to write entrance tests in order to enter the senior high school of their choice.

However, the school where this study took place (and where I teach) is a combined junior high and senior high school. Although learners are expected to maintain high academic standards, and may not be allowed to advance to senior high school if they fail too many subjects, there is not the pressure of high school entrance tests found in many junior high schools, and usually only one or two learners per year may be asked to find another high school.

At this school, learners usually start writing short compositions in the Junior 1 phase. Most topics concern personal introductions, as well as narratives based on short stories such as *Aesop's Fables*. In Junior 2, learners may be required to write about someone whom they admire, their favorite seasons, and what they did in the summer vacation. Some teachers may also introduce them to the importance of structuring their essays coherently. Since I started teaching at the school seven years ago, I have taught mostly Junior 2 students.

In my classes, I like to show the learners the different parts of a given composition. For example, a narrative usually starts with an introduction, followed by a meeting or parting.

This meeting or parting then causes a problem to occur. A large part of the story is then focused on finding a solution to this problem. The solution may or may not be favorable for the protagonist. In a longer narrative, an unfavorable solution will subsequently cause a new set of problems which then need to be solved.

1.3.3 Same skills, different structure

In the case of Japanese learners, there is evidence to suggest that skilled writers in one language are able to transfer these skills to writing in the target language (Sasaki 2000). On the other hand, since the approach and structure of Japanese texts are different from those of English texts, learners need to be made aware of how English texts are structured. This section deals with these factors in more detail.

Regarding the first point, it is clear that the ability to structure a text is an important skill in writing. Sasaki's study (2000) contrasted 'novice' writers with 'experts'. It showed how skills in the L1 can be transferred to the L2. The findings showed that writers who were better skilled in their L1 tended to plan their compositions more thoroughly than unskilled writers, and they were more concerned with the overall structure of the text than with sentence-level errors.

While the skills involved may be similar, there are, however, differences between Japanese and English texts in the kind of information they contain, and the way in which it is presented and structured. Firstly, while English texts place much of the onus on the speaker/writer to make his/her meaning clear, in Japanese it is the listener/reader who must make sense of what is being communicated. When developing an extended argument in English, the reader needs to be given signposts and signals as to what is being argued, and to where the speaker is in the course of the argument. While the western persuasive essay tends to focus on a single topic, with each aspect clearly indicated, the Japanese *ki-*

sho – ten – ketsu pattern of argument focuses on a number of topics. In this pattern of argument, evidence and supporting statements are cited first, followed by the topic sentence.

This is a pervasive form of essay writing, consisting of an introduction (*ki*) followed by the development of the introductory theme and loosely analogous subthemes (*sho* and *ten*) and a conclusion (*ketsu*) in which the essay makes its main point (Kimball 1996:58). There may, however, be leaps in logic where it is assumed that the reader understands the intervening steps.

In English, however, the main topic sentence is mostly stated first, followed by supporting sentences and evidence. Furthermore, while in English the conclusion is reserved mainly for summing up the argument with no new points introduced, in Japanese, new points may be introduced in the concluding paragraph, leaving a non-Japanese reader feeling as if the text has ended somewhere in mid-air. Thus, English writing is often described as linear, whereas the Japanese writing form (and the Chinese form from which it is derived) is often seen as circular (Fujieda 2006).

As a result of these differences, when teaching English writing to Asian students the teacher must be aware of teaching the structure of the text, and showing learners how all steps in logic must be carefully spelled out so that the reader can follow the writer's argument. However, as Matsubara (2001) points out, many Japanese teachers are themselves not aware of these differences, making it very difficult for them to teach the skills of English writing to their students.

1.3.4 Attention to audience

One difficulty often faced when assessing compositions is that while the grammar may be used correctly, the use of language and register is inappropriate or stylistically clumsy.

However, these problems with style and register are not given due attention, one of the reasons being that most writing in classrooms is only seen by the teacher, who then focuses primarily on grammar, vocabulary and spelling errors. Often the teacher has a sense that there is something wrong with the writing, but is unable to pinpoint exactly why this is. In these cases, the problem often lies in the fact that the audience has not been properly defined.

Coulthard (1994) states that there are in fact two types of reader: the real reader and the imaginary reader. When writing, it is impossible to know precisely who will be reading the text. It is therefore important to try to imagine who will be reading it. In so doing, one can form a better idea of what to assume this imaginary reader already knows, and what to describe in more detail (Coulthard, 1994). Elementary writers will often write statements of the obvious, such as 'the trees are green'. Nevertheless, one needs to be careful how one addresses such sentences. For example, the sentence 'Tokyo is in Japan' (my own example), when written for a Japanese audience may seem totally redundant, and possibly absurd. However, there may well be an imagined reader (such as a young elementary school student in an African, American or European country) for whom such a sentence would be new information. If one is to improve students' writing, one needs to work on giving them the opportunity to read each other's writing, or to have their writing read by a wider audience. As Coulthard (1994) notes, some texts fail because the author neglects to maintain a consistent imagined reader.

While attention to audience is not the main focus of this study, it is still a very important factor to consider when assessing texts, and can, if not properly developed, have a significant impact on the cohesion of a text.

From the point of view of cohesion, attention to audience is about deciding which information needs elaboration and which may be more concisely provided or left out altogether. Thus, there may well be stronger cohesive ties between sentences where ideas are elaborated on than between those containing less important information.

1.4 DEFINING THE RESEARCH PROBLEM

I am currently teaching Japanese junior high school students between the ages of 12 and 15 years (the equivalent of Grades 7–9 in South Africa) who, with the exception of a few returnees from English-speaking countries, only started studying English when they arrived at this school aged around 12 years of age. Given that students will eventually have to write essays to pass their university examinations, the main aim of teaching essay writing to these students is to introduce them to a variety of writing genres so that they will be more familiar with essay writing, and will be equipped with the necessary skills to tackle essays, not only in their university entrance tests but also at university and possibly in their places of work.

Although many teachers in Japan (particularly those teaching at private schools) are highly skilled in the teaching of writing, textbooks on essay writing tend to focus on the higher grades, namely, Senior 2 and Senior 3 (equivalent to South African Grades 11 and 12). The common assumption is that students need to learn from the 'bottom up', starting with words, building up to phrases and sentences and, finally, to extended texts. This means that there tends to be less emphasis on essay writing, particularly in the lower grades where knowledge and understanding of vocabulary and grammatical structures are limited. Moreover, much of the research that has been conducted has focused (as so much academic research does) on university students.

The question that arises is what is happening in texts written by students in the elementary to pre-intermediate phases of language learning? How are they constructing sentences, and what does their use of cohesive devices reveal about their ability to write a successful text? Is it the number of cohesive devices used that is important, or is it the way in which these devices are used which enhances the quality of the text?

One way of addressing these questions is to look not only at the writing of the students themselves, but also at the textbook which they are using. The textbook (which will be described in more detail in Chapter 3) provides a number of reading passages and writing exercises that are important models for the teaching and learning of writing. This leads to the question of how the learners' use of cohesive devices compares with the text models that they use.

1.5 RESEARCH AIMS

The ultimate aim of this study was to determine as accurately as possible what cohesive devices were being used by beginner to lower intermediate learners, and how this use compared with the use of devices in the prescribed textbook. Establishing which cohesive devices were being used by learners at various specified levels would make it easier to devise ways of teaching these cohesive elements at these particular levels.

1.6 ASSUMPTIONS AND RATIONALE

Since each reading passage in the textbook is aimed at a Grade 7 level of proficiency (which is, in fact, middle to upper elementary level), it is reasonable to assume that the cohesive devices which occur in these reading passages are those which these students could be expected to comprehend and use in their own compositions. The textbook therefore serves as a model and a yardstick by which to measure the learners' grasp of composition in general, and use of cohesive devices in particular.

Given this premise, the rationale behind analyzing reading passages from the textbook lay in the fact that these reading passages were written in such a way as to review the grammar and vocabulary that had already been covered in the textbook, as well as to introduce new vocabulary and functional expressions presumably set at grade level. The reading sections can therefore be regarded as a good indicator of the kind of language that students at this level could be expected to use.

It was hypothesized that the learners' compositions would offer a rich source of data relating not only to cohesion but also to other elements of composition. Moreover, I believed that closer qualitative analysis of some of these compositions would allow for a deeper insight into the kinds of cohesive devices used by these learners, as well as what cohesive elements might be appropriate at this particular level of language proficiency.

Furthermore, written essays are one way in which learners are able to display and apply their understanding and acquisition of the language. Since writing is a planned process, it allows the writer time to formulate what he or she wants to say, and how he or she wants to say it. Not only can it indicate levels of proficiency, but the errors often reveal areas of language use which the teacher may need to revisit. Essays therefore provide the teacher with excellent insight into a learner's language acquisition and proficiency. In a class of 45 to 47 students, where effective monitoring is often very difficult, composition offers an important view of learners' abilities.

1.7 THE RESEARCH QUESTIONS

The rationale leads to the main research question:

Main research question

- Assuming that the reading passages in an English textbook for second-year Japanese junior high school learners provide a model for the type and quality of writing that may be expected at this language level (namely, upper elementary), how does the use of cohesive devices reflected in these compare to that of the writing by Grade 7 Japanese learners?

The question is broken down into the following sub-questions:

1. What cohesive devices are found in the textbook readings (TBRs), and what cohesive devices are found in the student texts (STs)?
2. Is there a difference in frequency in the use of cohesive devices in TBRs and in STs?
3. What patterns of reiteration occur in the TBRs?
4. How do the patterns of reiteration in the STs compare to those in the TBRs?
5. Which cohesive devices in the STs are used inappropriately or inaccurately?
6. What are the reasons for these errors?

To this end, the study analyzed a sample of 36 essays from a class of 180 learners in Junior 2 (equivalent to Grade 8 in South Africa) at a private school in Kanagawa, Japan. The essays comprised book reviews written under test conditions as part of the end-of-term test. Eleven TBRs that had been read up to that point in the school year were also analyzed. Two types of analysis were carried out. Firstly, the quantitative analysis manually identified and coded cohesive devices in the texts according to the taxonomy laid out in Halliday and

Hasan (1976). The coded cohesive devices were subsequently collated and converted to percentages for statistical analysis. In the qualitative study, two TBRs and six STs, two graded by me and four other raters as successful, two graded/assessed as average and two essays graded as unsuccessful, were analyzed according to the patterns of repetition adapted from Hoey (1991). The STs were analyzed to determine the extent to which they contained the features found in the textbook readings. Finally, the cohesive errors found in the student texts were analyzed and the possible reasons for these errors were examined.

1.8 THESIS STATEMENT AND DELINEATIONS

Given the questions posed in 1.7 above, it follows that this study will be discussing the following thesis:

It is possible to gain a clearer picture of the strengths and weaknesses of Japanese Grade 2 junior high school learners' compositions by comparing the cohesive devices found in them to those found the reading sections of the textbook they are using.

1.8.1 Delineations

It should be noted that the picture that the study hopes to gain is confined to the use of cohesive devices. Of course, there are other important criteria for good writing. The question in this study is how the use of cohesive devices helps or hinders the success of a text.

The compositions used here are texts written by learners, most of whom have only been studying English for a year and three months, and whose language proficiency may be deemed elementary or upper elementary. The exact limitations of the learners' language abilities are detailed in Chapter 3 (§3.5).

1.8.2 Definition of terms

Text quality: This term covers the aspects defined in §1.2.2 above, namely, structure, content, flow of language and accuracy.

Cohesive devices: In this study this term refers to those grammatical and lexical devices that link sentences to one another, thus contributing to the creation of a complete, coherent text. The categories of grammatical cohesion identified by Halliday and Hasan (1976) (see §1.3.1 above) were used in the quantitative study. These categories, along with the types and patterns of repetition presented in Hoey (1991), inform this part of the study and are discussed in more detail in Chapter 2.

Japanese junior high schools refer to schools which accommodate learners from the equivalent of South African Grade 7 to Grade 9.

Learners: The term 'learners' refers to the learners who attend the school at which the study took place. This school is a Catholic-run boys' school near Tokyo, Japan.

Compositions and **essays:** These are extended texts such as narratives, letters and essays that are written by students. They may cover different genres such as narrative, descriptive and persuasive writing. While the term 'composition' may cover a broader range of genre than 'essay', the two terms are used interchangeably in this study.

1.9 SUMMARY

The chapter started by providing some background to this study of cohesion by highlighting the prominence of grammar translation in the Japanese English high school curriculum, and briefly tracing the move toward more communicative approaches. It was noted how, while the focus has been more on speaking and listening than on reading and writing, the

use of composition in university English entrance tests has meant that teachers in Japanese high schools now have to focus on teaching composition.

My own teaching situation was briefly described, and the possibility of teaching composition in the lower grades of high school was suggested in order to take some of the pressure off the higher grades

This chapter mentioned the different grammatical and lexical cohesive devices as identified by Halliday and Hasan (1976). These are discussed in more detail in Chapter 3.

The chapter then presented the main aims of the research, the research questions and the rationale behind using the learners' compositions and the prescribed textbook as the main source of data, before defining some important terms that will be used throughout the study.

The remainder of the dissertation is structured as follows:

Chapter 2 provides a review of the literature concerning essay writing in general and cohesion in particular. It looks at studies of cohesion – those that made use of frameworks similar to those used in this study, and those carried out at Japanese universities and schools.

Chapter 3 discusses the approach, methods and methodology of the study. It describes the aims of the study, the framework of the study, the sampling methods used and how the data were analyzed; that is, both quantitatively and qualitatively. This chapter also explains the ethical considerations and how these were addressed.

Chapter 4 presents and discusses the results of both the quantitative and the qualitative aspects of the analysis.

Chapter 5 discusses the implications of the findings. It explains the implications of these findings for teaching, as well as the limitations of the study, finally offering recommendations and suggestions for further research.

Chapter 2: Review of the Literature

2.1 INTRODUCTION

This chapter presents a review of the literature relevant to this study. It begins by explaining briefly the nature and importance of cohesion. This is followed by a brief discussion of some of the most important works on the subject, with particular emphasis on Halliday and Hasan (1976) and Hoey (1991), both of whose work form the frameworks for this study.

The second part of this chapter examines some of the more recent studies that have used as frameworks the work of either Halliday and Hasan, or Hoey, or both. This chapter considers what these studies reveal about cohesion in writing, as well as their relevance to this particular study.

The main goal of this study was to compare the use of cohesive devices in writing by students in a Japanese junior high school with their use in the prescribed textbook. Using the textbook readings as a model, the study aimed to ascertain and better understand the students' strengths and weaknesses in terms of their use of cohesive devices. To this end, some background on, and some of the key concepts and terms used in, cohesion will be discussed. In addition, the key frameworks used for this study will be examined. Some of the studies that have taken place using these frameworks, and the extent to which they address the questions posed by this study, are then discussed.

Studies of cohesion may be broadly categorized according to (i) their aims, (ii) their frameworks, (iii) their approaches, participants and methodologies, and (iv) their findings.

Regarding the aims of studies of cohesion, three main issues have emerged:

1. The correlation between cohesion and successful writing
2. Precisely which cohesive devices students use in their writing
3. Cohesive errors and the influence of the L1 on second or foreign language compositions.

The studies discussed below all used either Halliday and Hasan's (1976) framework, or Hoey's (1991) framework, or a combination of the two. Studies using the former tend to be quantitative, while those using the latter tend to be qualitative. Most studies approach the study of cohesion by comparing texts that differ in one way or another, be they native or non-native texts, advanced or elementary, or graded high or low. For example, Abadiano's (1995) and Castro's (2004) studies compared texts from different levels of proficiency, while other studies used a model text created by an adult with native proficiency (Hinkel 2001; Al Khafafi 2005). Still other studies compared different genres of texts, such as comparisons between narrative and argumentative writing (Chen 2007).

Participants in most studies were university students, who were clearly experienced language learners (Chen 2007). One study (Abadiano 1995) used elementary school students, but these were in a native-speaking setting. Two Japanese studies, namely, Nakao (2009) and Okuda (2012), used high school students. This was closer to the setting in the present study, but the learners in Nakao's and Okuda's studies were still clearly more advanced than the learners in the present study. There were also studies which analyzed only one type of text, for example texts by Arab speakers (McGee 2009), or English learners in Nigeria (Olateju 2006). The findings of the studies discussed in this literature section, can be summarized as follows:

- Most studies show that the frequency of cohesive devices is not necessarily a significant factor in the quality of the texts.
- They show that in the foreign language texts, some individual devices were used more or less frequently than others.
- They show that the ways in which devices were used differed between successful and unsuccessful texts. For example, the use of patterns of repetition tends to differ between more successful or less successful texts. Moreover, successful texts tend to elaborate more, allowing for more bonded sentences.
- Studies highlight some of the errors and problems in L2 scripts, and show how and why these errors occur. More specifically, they reveal how the L1 can influence the use of cohesive devices.

The core argument I shall be making in this chapter is that, while previous studies offer valuable insight into the use of cohesion by EFL and ESL learners of English, their findings are not wholly applicable my study, since none of them involved learners at an elementary level of proficiency. The participants in the studies mentioned above were university students, senior high school students, or young learners in an L1 context. An examination of the literature will expose some of the gaps in our knowledge of the use of cohesion by elementary-level learners of English in general.

The main question, therefore, is what precisely the present study can contribute to our understanding of cohesion in EFL texts. There are two answers to this question. Firstly, because of the many similarities between this and previous studies in the frameworks and approaches used, this study may serve to either strengthen or refute previous findings. Furthermore, given the fact that this study focuses on younger learners with barely a year of English language learning behind them, any new findings can only increase our

understanding of some of the problems faced by elementary learners, and offer some clear data with which these problems may be addressed.

Having outlined the main argument of this chapter, I turn now to providing some context, starting with a brief discussion of cohesion versus coherence.

2.2 COHESION AND COHERENCE

The first question that needs to be examined is how one distinguishes between cohesion and coherence. The difference between the two can be illustrated by the following example:

Example 2.1

Peel four potatoes.

Put them in some boiling water.

In Example 2.1, the word 'them' presupposes something that comes before it, namely, the potatoes. This reference to something in a previous sentence creates a link between the two sentences, and constitutes cohesion between them.

Coherence, on the other hand, is concerned with how the meaning of the text hangs together, and how the text is interpreted as a whole. This meaning, according to the proponents of schema theory, is derived not simply from the text itself, but from the knowledge and expectations that the reader brings to the text (Carrell 1982:482). In the above example, the reader may know that the words *peel*, *put* and *boiling water* have something to do with cooking, and that recipes are often written in the imperative. He or she may then match this schema to the text, finding that it matches all the criteria for a recipe. Of course, the process is taking place instantaneously and unconsciously. This example also shows how features of a text include the way in which the sentences are arranged, the rhetorical features of the text, and how the text addresses the reader's expectations.

This distinction between the quantifiable links in cohesion and the more subjective schemata that different people may bring to a text can be seen in Hoey (1991). He sees cohesion as the objective property of the text – that which can be statistically counted – while coherence, because it can be evaluated only according to the reader's judgment, which varies from reader to reader, is to a large degree subjective (Hoey 1991:11–12). In the example of the potatoes (Ex. 2.1 above), it is possible to count the cohesive ties (*potatoes, them*), but whether the above is in fact recognized as part of a recipe or not may depend on the knowledge and experience of the reader.

These viewpoints suggest two important points: Firstly, coherence refers more to the overall structure of the text and the moves that are appropriate to whichever genre the text belongs, while cohesion can be seen as the individual semantic ties that bind the sentences together. Secondly, cohesion can be quantifiably measured, whereas coherence is more dependent on the reader's interaction with the text (Castro 2004:216).

Having briefly examined the distinction between cohesion and coherence, the next question that arises is what precisely the nature of cohesion is, and what precisely is its relationship to coherence? Is cohesion necessary for coherence, or is the opposite true? In order to answer these questions, some historical context is required.

For Halliday and Hasan (1976), cohesion brought about coherence. However, researchers began to question this central role of cohesion. Carrell (1982) was particularly critical of this claim. She saw the schema which the reader or listener applies to the text as more important in reading and listening comprehension. Other researchers attempted to show that it was possible to create coherent texts without the use of cohesive devices. Perhaps the most famous example appeared in Widdowson (1978 cited in Fulcher 1989:148):

Example 2.2

A: That's the telephone
B: I'm in the bath
A: OK

In this example, the exchange is perfectly coherent, but does not contain any cohesive links between each sentence. It is thus claimed that the coherence lies in the schematic understanding of the text rather than in the presence of cohesive devices.

This example suggests that cohesion and the classification of cohesive devices, while useful in themselves, are not central to coherence. However, these warrant further examination. Fulcher (1989) points out that this exchange, while arguably a text, also takes place across conversation turns, which usually involve ellipsis. My own view of the above is that the cohesion actually lies in the parts that have been elided (see Ex. 2.3 below). If one were to add the cohesive devices, the exchange might read as follows:

(Here, the boxed parts of the text indicate elision, while bracketed words show the cohesive items.)

Example 2.3

A: That's the (telephone) ringing.

B: I can't answer (it,) because I'm in the bath.

A: OK. I'll get (it.)

The diagram consists of three lines of text. In the first line, 'ringing' is enclosed in a box. A vertical line descends from the bottom of this box to the top of another box in the second line. In the second line, 'I can't answer (it,) because' is enclosed in a box. A diagonal line descends from the bottom of this box to the top of a third box in the third line. In the third line, 'I'll get (it.)' is enclosed in a box.

As can be seen, simply because the cohesive items were not mentioned does not necessarily mean that they are non-existent.

Conversely, examples of texts containing cohesive devices but no coherence were also cited. Kolln (1999) mentions one example to show how some researchers believed that, since cohesion refers to the links between sentences, it was possible to have a cohesive text that was not coherent:

Example 2.4

My computer is on *my desk*.

My desk is made of *oak*.

Tall oaks from little acorns grow. (Kolln 1999:112)

In Example 2.4 above, certain words (*desk, oak*) are repeated, and yet the text makes no sense. However, Kolln (1991) questions whether Example 2.4 above is a text at all, since it has no unity or meaning as a whole. It is, in fact, simply a collection of sentences connected by the first and last words of each sentence. It falls under no particular genre, performs no discernible function, and contains none of the modes of language found in a true text.

As Example 2.4 shows, both cohesion and coherence are essential if a piece of writing is to be defined as a text. In this example, although the word *oak* is repeated, the second *oak* has no relation to the first; it refers to oak trees in general, rather than to the oak desk. Moreover, apart from the word *oaks*, there is no other reference to or added information about the computer or desk mentioned in the previous sentences. The reader may deduce that the first two sentences constitute the premises of some kind of logical argument, or perhaps the beginnings of a description of the writer's office equipment. However, the third sentence does not contain any direct reference to the previous sentences, nor does it add any new information. It therefore does not match the criteria for any known genre of text.

Halliday and Hasan's (1976) taxonomy gave rise to the question of whether the frequency of cohesive devices had any relationship to the quality of the text. In one of the earlier studies that addressed this question, Witte and Faigley (1981) compared the essays written by University of Texas college freshmen. Out of a total of 90 English essays that had been holistically graded by two different raters, they analyzed five highly graded essays and five low-rated essays using error analysis, according to syntactical features, and the 'number and types of cohesive ties' (Witte and Faigley 1981:195). The findings of their analysis suggested that cohesion was not the main prerequisite for coherence, despite a greater frequency of cohesive devices among the higher ranked essays (Witte and Faigley 1981). They believed that factors beyond the text, such as audience, point of view and the reader's knowledge of the subject played a more crucial role (1981:199). For them, collocation, one of the devices identified by Halliday and Hasan (1976), could be seen as the best indication of writing ability. Indeed, they reported that higher-rated essays tended to feature more collocation than low-rated essays (Witte and Faigley 1981). Witte and Faigley's study also raised the question of how best to account for the important role that schemata play in coherence.

The disproportionately greater attention paid to grammatical devices over lexical devices was another problem in Halliday and Hasan's model (1976). Indeed, as Hoey (1991:9) pointed out, the number of pages devoted to all five lexical devices (about 20) is less than half of the 50 pages devoted to substitution alone. Moreover, while grammatical devices were precisely and thoroughly classified, devices contained areas that were, even in the words of Halliday and Hasan, 'problematic' (1976:284), since they left room for interpretation in coding.

Hasan (1984:202) later addressed these deficiencies by modifying her classification of lexical devices as follows:

A: General

- i: repetition – leave, leaving, left
- ii: synonymy – leave, depart
- iii: antonymy – leave, arrive
- iv: hyponymy – travel, leave
(including co-hyponyms, leave, arrive)
- v: meronymy – hand, finger
(including co-meronyms finger, thumb)

B: Instantial

- i: equivalence – the *sailor* was their *daddy*; you be the *patient*, I'll be the *doctor*
- ii: naming – the *dog* was called *Toto*; they named the *dog* *Fluffy*
- iii: semblance – the *deck* was like a *pool*; all my *pleasures* are like *yesterdays*

Another important addition by Hasan (1984) was her focus on lexical chains, which are words or ideas that are repeated successively through different parts of the text. She identified two types of lexical chains, namely, identity chains and similarity chains. Identity chains are those where the words have the same referent. For example, in the first and second sentences of this paragraph, the following lexical chain was formed:

Example 2.5

Hasan

She

Similarity chains occur when the same word is repeated, even if the repeated word has a different referent. For example, if the word *run* is repeated, it forms a similarity chain, even if two different beings did the running. While this present study did not focus so much on

similarity chains, identity chains formed an important part of the qualitative analysis (see §4.5).

While Halliday and Hasan (1976) classify cohesion into different kinds of cohesion, Hoey (1991) sees cohesion largely in terms of the way in which different words or ideas are repeated or reiterated throughout the text. Furthermore, Halliday and Hasan (1976) focus largely on grammatical cohesion, while Hoey is more concerned with lexis.

Finally, the dichotomy between quantifiable cohesion and more qualitative coherence in Halliday and Hasan (1976) is largely removed in Hoey's more precise and objective way of describing the relationships between sentences. In Hoey's work, the term *repetition* is used in a much wider sense. He notes that the main function of repetition is to 'allow a speaker to say something again so that something new may be added' (1991:53). This is an important point, since it implies that if repetition does not allow for the addition of new information, it is probably redundant. Moreover, while Halliday and Hasan (1976) include such grammatical items as conjunction and substitution in their taxonomy, these are not included in Hoey's (1991) study.

Using Winter's clause relations (1974 and 1979, cited in Hoey 1991:16–18) and Phillips's collocation links and clusters (1985, cited in Hoey 1991:21–25) as his point of departure, Hoey makes a case for a focus on the reiteration of lexical items, and the clustering of repetition in sentences (Hoey 1991:20). He cites Phillips's (1985 in Hoey 1991:23) work in the way in which different collocations within a given chapter are connected to one another. Where networks of words in different chapters resemble one another, the chapters themselves could be said to have a closer relationship. Indeed, it was Phillips (1985 in Hoey 1991:22) who first demonstrated that sentences that are linked by three or more reiterated ideas may form a coherent bond.

However, Hoey (1991) took this further by showing how these bonds occurred not just between sentences placed a few sentences apart, but across vast stretches of text, and even between different volumes of text. What Hoey (1991) therefore shows is that cohesion is indeed an important part of coherence, and that it is the patterns of reiteration that occur between sentences rather than the frequency of different cohesive devices that are crucial in both the cohesion and the coherence of a text.

Hoey's (1991) model includes a number of different types of repetition, namely, simple repetition, complex repetition, simple paraphrase, complex paraphrase, as well as superordinate, hyponymic and co-reference repetition. Since my study was more concerned with patterns of repetition than with specific types of reiteration (which are covered in Halliday and Hasan's [1976] taxonomy), these specific types of reiteration were not recorded during data analysis. However, in order for the reader to gain some understanding of how and why certain lexical items came to be classified as repetition, as well as how they are related to Halliday and Hasan's (1976) taxonomy, they are described in some detail in Chapter 3 (§3.5.2) below.

2.3 THE KNOWN-NEW CONTRACT

Kolln's (1999) known-new contract explains how a well-constructed sentence usually contains known information in the first part, and new information in the verb and complement part of the sentence. The known-new pattern helps to break the monotony when known information is repeated. Furthermore, the reader or listener knows where to find the new information, while words that contain new information are emphasized. Finally, placing this new information in the latter part of the sentence can help with rhythm and intonation.

While the known-new contract is not a central focus in this study, it does help to explain why some texts may be stylistically inadequate. This is because the writer of a weak text may place the known information at the end of the sentence. For example, the following type of sentence frequently appears in the writing of the learners whom I teach:

I like summer. There are watermelons in summer. I always visit my grandparents in summer.

As can be seen in the above example, the known information in the second and third sentences is placed at the end of the sentence, making the text seem somewhat repetitive. Moreover, understanding how the known-new contract works also means that the writer is better able to either eliminate redundantly repetitive patterns, or find other ways of reiterating them, such as by the use of synonyms or superordinates.

Having provided a brief overview of cohesion and its relationship to coherence, as well as tracing the work of Halliday and Hasan (1976) and Hoey (1991), the remainder of this chapter focuses on some of the studies that have focused on cohesion in writing.

2.4 STUDIES IN COHESION

Ever since Halliday and Hasan's (1976) definitive taxonomy of cohesive devices, a large number of studies have been carried out on the subject of cohesion in writing. As noted in §2.1 above, these studies may be broadly categorized according to their aims and the types of issues they address. These issues include the following.

- the relationship between cohesion and successful writing
- precisely which cohesive devices students use in their writing
- whether cohesive devices are used correctly or incorrectly.

These issues, as well as their relevance to the present study, are discussed below.

2.4.1 Cohesion and language proficiency

On the topic of the relationship between the frequency of cohesive devices and successful writing, Witte and Faigley's study (1981), which is discussed in §2.3 above, has proved to be something of the exception to the rule, since it found a significant relationship between cohesion and text quality. Many studies have found no significant relationship between the frequency of cohesive devices and the quality of writing. Abadiano's (1995) comparison of writing by 24 Grade 6 elementary school students focused on three distinct varieties of English ('mainstream', Appalachian and African American), none of which showed any significant differences in the use of cohesive devices. (Other findings from this study are discussed in §2.5.2 below.) Of course, it may be argued that the writing of learners in what was essentially a first-language setting could have no relevance to the writing of learners in an Asian environment. However, as will soon be shown, the nationalities of the learners did not seem to be a factor in this regard.

Studies among Asian learners were also unable to find any direct relationship between cohesion and the success or otherwise of a text. Using, by her own account, a relatively homogenous group of university students, Castro's (2004) study of essays by 30 Filipino college students revealed no significant differences in the number of grammatical cohesive devices used by the two groups, while lexical devices and the use of synonyms were the most common ways of achieving cohesion (Castro 2004:222).

While the studies above focused on grammatical cohesive devices, Mojica's (2006) study focused on lexical devices. Her study also found no relationship between the frequency of devices and the success of a text. Mojica (2006) investigated 30 graduate students of advanced academic writing at a university in the Philippines. Her study addressed three questions, namely, the types of lexical cohesion that were employed by ESL students in advanced writing courses; 'how lexical items cohere with a preceding occurrence of the

same item', and finally, the question of what students' holistic scores in their essays may suggest (Mojica 2006:109). Participants were divided into two groups, namely, a multidisciplinary group (A) and an English learning group (B). The advanced writing course taken by group A was meant only for graduate students in various fields who had failed to pass the essay part of their admissions tests. The two groups were presumed to be at different levels of proficiency in English. The study used Halliday and Hasan's (1976) taxonomy of cohesive devices. In this taxonomy, four of the five lexical devices (namely, repetition, synonyms, superordinates and general words) are all seen as types of reiteration (Halliday and Hasan 1976:278).

The first part of the conclusion of each paper produced in Mojica's study (2006) was examined for lexical cohesion by two professors. Scores were awarded on a scale of 1 to 4, with 4 being above average, and 1 being not cohesive. The texts were analyzed using four types of lexical ties, namely, Group I: repetition and synonyms; Group II: superordinates and hyponyms (for example, *animal* is a superordinate of *dog*, while *dog* is a hyponym of *animal*); Group III: related words; and Group IV: text-structuring words such as *agenda*, *advantage*, *problem*, *reason* – namely, words which in Halliday and Hasan's taxonomy would be 'general words', and which have a more specific referent either before or after them. Student group B was actually behind group A in the use of synonyms, superordinates and hyponyms. This was unexpected, as language students were expected to be more familiar with hyponyms than the multidisciplinary group (Mojica 2006:114). In fact, it was only in their ability to use text-structuring words that group B performed better than group A.

The studies of Castro (2004) and Mojica (2006) bear some resemblance to the present study in that they compare successful texts with unsuccessful texts. Furthermore, Castro acknowledges that her sample is largely a homogenous one, drawn from people from very similar backgrounds. Castro's (2004:222) study suggests that the reason for the lack of

differences in grammatical cohesion is the homogenous characteristics of the sample group, and that writers who share common socio-cultural backgrounds tend to draw on shared experiences and schemata in the construction of their essays. This view lends support to Carrel (1982), and raises some questions in the present study, where, given their gender, nationality, high level of academic achievement and overall motivation to study, participants were also drawn from a homogenous background. It was therefore important to explain the use of cohesive devices, not simply in statistical terms but also within the context in which the essays were written, and the knowledge and schemata from which students drew in the writing of their essays.

Castro (2004:222) concludes that one also needs to look at the sociolinguistic context in which a study takes place, since this can determine the characteristics of the semantic item. Writing involves an interaction of schemata, which is the background knowledge one carries of a particular topic or item. In homogenous groups such as the one in her study, the subjects all carried with them similar frames of reference and schemata. This explains the 'minimal differences' between the groups (Castro 2004:222). As her study suggests, statistical analysis alone is often inadequate in accounting for the use of cohesive devices, particularly in homogenous samples.

The same can be said for the present study, since all the learners are Japanese, have come through the Japanese school system and are very similar in terms of socio-economic status. However, as noted in §2.1 above, these studies differ from my study, firstly in the use of university students as participants, and secondly because learners in the Philippines are exposed to English from a much earlier age than in Japan. A final difference between Mojica's study and the present study lies in the fact that her 'advanced' group was also a group of L2 learners, rather than native-level speakers. The present study, on the other

hand, compared the writings of Japanese learners to a model text written by native-level speakers.

The lack of a strong relationship between text quality and the frequency of cohesive devices is also true of different genres, according to Chen (2007), who investigated lexical cohesion in Chinese college EFL students' writing. Apart from investigating whether certain types of lexical cohesion dominated EFL writing, he also wanted to determine the extent to which lexical cohesion was affected by students' ability, or the type of text. For example, does a narrative text have an effect on the kind of cohesion used? It is precisely this kind of question that was addressed in the present study.

In his investigation, Chen (2007) used the writing of 30 English majors from a university in China. Subjects were asked to write two texts, one narrative and one argumentative, with a one-month interval between the two tasks in order to 'avoid inter-genre influences' (2007:47). The samples were then analyzed using Halliday and Hasan's (1976) taxonomy because of its analytical quality, and careful explanation of lexical cohesive ties (Chen 2007:47).

Chen's findings did not show any significant relationship between language proficiency and lexical cohesion (Chen 2007). This supports the findings of Castro (2004) and Abadiano (1995). Accordingly, a 'higher mean frequency of total lexical cohesive devices in narration' (Chen 2007:49) was found, and a stronger correlation between lexical cohesion and text type. There was also more collocation in narration, but more reiteration in argumentation. However, 'the difference in the mean frequency of total lexical cohesive devices between narration and argumentation is not significant' (Chen 2007:49).

The weak relationship between language proficiency and lexical cohesion in Chen's study is explained firstly in terms of the fact that the difference in language proficiency of the two

groups may not be as great as is implied by their level at university. Another reason could have been that, although passive vocabulary knowledge had increased, the students still needed practice in activating various words (Chen 2007:52). As in Castro (2004) and Mojica (2006), the use of university learners as participants makes it difficult to apply these findings to the present study. Moreover, in Chen's study, the texts being compared were both non-native texts, unlike my study which used an English textbook as a model against which the learners' writing was compared.

Later studies in Japanese schools by Nakao (2009) and Okuda (2012) also found no significant relationship between the frequency of cohesive devices and the success of a given text. These studies are described in more detail in §2.5.5 below. The question that remains is whether my study would support or refute the findings above, especially in the case of learners at an elementary level of English proficiency.

While none of the findings of the studies cited above revealed any significant relationship between the frequency of cohesion and the success of a text, further investigation revealed more precisely whether there were any significant differences in the types of cohesive device used. These findings, along with those of other studies, are discussed in the following sections.

2.4.2 Types of cohesive devices in EFL writing

The lack of a significant relationship between successful writing and the number or frequency of cohesive devices that emerged from many studies gave rise to the second question, namely, which devices were to be found in the compositions of EFL and ESL learners.

The use of different cohesive devices can be seen in Abadiano's (1995) study, which showed that most essays used lexical cohesion, particularly repetition, as well as reference

and conjunction (Abadiano 1995:307), while substitution and ellipsis were 'hardly ever used' (1995:307). She also observed that demonstrative reference was highest in the mainstream sample. Abadiano noted, firstly, that all the groups used all five devices (namely, reference, substitution, ellipsis, conjunction and lexical items), and that repetition and conjunctions were the most frequently used. The mean percentage of occurrence of each cohesive link 'varied from child to child, piece to piece and group to group' (1995:307). Nevertheless, there was a significant variation in the different types of cohesion found. Thus, in this study, she found that it was not so much how many cohesive devices were used as which types, and where. This is why the qualitative analysis in the present study was so important.

One study which examined how cohesive devices were used was that of Sardinha (1997), who analyzed a corpus of 300 texts covering three genres, namely, research articles, business reports, and encyclopedia articles. This study found that texts signaled shifts in topic through 'breaks in lexically cohesive clusters' (Sardinha 1997:10). These findings suggested that the role of discourse markers (such as *however*, or *in conclusion*) was only one part of the way in which sections of a text are divided. Sardinha's work has particular relevance to my study when examining the patterns of repetition in the text. His study is significant given the large body of text that was used, which can only enhance its reliability. However, since it focuses on quite sophisticated texts by native writers, the question still remains as to the extent to which the implications of these findings may also be true for elementary-level learners of English.

One of the aims of the present study is to determine how the student texts and the textbook readings differ in their use of grammatical and lexical devices. With this in mind, we now take a closer look at the use of demonstratives, repetition and synonyms.

2.4.2.1 Demonstratives, conjunctions and sentence transition

Some studies (Abadiano 1995; Hinkel 2001; Liu and Qi 2010) have noted the difference in the use of demonstratives, conjunctions and sentence transitions in EFL writing.

In the present study, the learners still had a very limited number of options for employing sentence transitions. The question was what these were, and how they intended to use them.

While most writers use various features of cohesion in their texts, certain elements increase with the sophistication of their writing (Kolln 1999:96). For example, younger or elementary learners tend to use 'then' to signal time in the lower grades, whereas higher grade students will use 'first of all', 'next', 'for one thing', 'meanwhile', 'all in all' and 'finally' (Kolln, 1999). Abadiano's study, as noted in §2.5.2 above, found that conjunctions were, along with repetition, one of the most commonly used cohesive items in both the 'mainstream' and non-mainstream groups. Hinkel (2001) notes in his analysis of 897 academic texts by Japanese, Korean, Indonesian and Arabic university-level learners of English that, regardless of the L1, non-native speakers (NNSs) tended to use significantly more sentence transitions (for example, *then*, *first*, etc) and demonstrative pronouns (*this*, *that*) than native speakers (NSs). Moreover, even NNSs with fairly advanced proficiency tend to rely more heavily on grammatical cohesive devices than on lexical devices. For example, Arabic speakers made greater use of coordinating conjunctions such as *and* and *or*, while other NNS texts contained a high frequency of the demonstrative *this*.

The use of more sentence transitions and conjunctions can also be seen in Hinkel's (2001) results, which showed that speakers of Korean and Japanese used coordinating conjunctions at similar rates to those found in NS essays (Hinkel 2001:120), Indonesian students' writing contained fewer such markers, while Arabic students' writing contained

significantly more. Moreover, the NNSs' use of sentence transitions was significantly higher than that of the NS group (Hinkel 2001:122).

Hinkel suggests that this could be because NNSs tend to rely heavily on sentence transitions (*therefore, in addition, however* or *secondly*) to make their texts cohesive (2001:123). Hinkel (2001:123) also suggests that one of the problems is that sentence transitions are often given more emphasis than is necessary, leading to their overuse in L2 texts.

Hinkel's findings raise the question of the extent to which L1 plays a part in learners' choices of cohesive device. His use of a large sample (n = 897) gives the study considerable weight in its applicability to the general population. However, since the participants were students in America, it hard to know whether these results would necessarily be found among elementary-level learners in an ESL or EFL environment.

The question regarding the present study, therefore, was which sentence transitions would be used by Japanese school learners at an elementary language level, as well as how these would be used. It also raised the question of whether, in the present study, lexical or grammatical devices would be more prevalent in the student texts or in the textbook readings.

2.4.2.2 Repetition

Several studies of college-level EFL writing have found repetition to be the most frequently used type of cohesive device (Castro 2004; Mojica 2006; Chen 2007). In Chen's study (2007), it was used three times more than all the other lexical devices (synonyms, collocation, superordinates and antonyms – in that order of frequency) combined. Chen (2007) also examined the mean frequency of each subtype of reiteration. Exact repetition

was the most common, followed by synonyms, superordinates and general words, with collocation used least.

While repetition is indeed an important device, overuse can have a negative impact on the quality of the composition. What is important is how repetition is used. McGee's examination of repetition and collocation referred to the use of what he called 'redundant repetition' (2009:213), as opposed to repetition that can and does enhance the effectiveness of a text.

A certain amount of exact repetition can be an effective way of showing cohesion. Indeed, it may even be necessary, as in the case of legal documents (Cook 2004:19). In fact, this is supported by the work of Hoey (1991). Repeating topic-related words helps to thread a text together at both the structural and the semantic level. Moreover, in Mojica's study (2006), the use of repetition was found to be an easy way to set up lexical ties and to create emphasis.

Some interesting insight into how direct repetition can affect a text can be found in Al-Khafafi's (2005) study. One objective of this study was to test the validity of two common assumptions about English and Arabic texts, namely, that:

- A. There is more repetition than variation in Arabic lexical repetition chains.
- B. There is more variation than repetition in English lexical cohesive chains (Al-Khafafi 2005:7–8).

The frequency of repetition in Arabic and English texts was investigated by Al-Khafafi's (2005) study, which analyzed in depth two argumentative texts, one in Arabic and one in English, belonging to the same genre. The texts were then analyzed for lexical cohesive chains. All cases of simple and complex repetition were noted. In simple repetition, a word is repeated directly, or with minor alteration. Complex repetition refers to words that are

repeated, but where the grammatical function changes, for example the verb *arrive* may be repeated as *arrival* in a subsequent sentence.

The initial findings showed a remarkable similarity between the two texts in the ratio of simple repetition to complex repetition, and found assumption A above to be valid while assumption B was not. This also raised the question of why Arabic is commonly believed to contain more and English less complex repetition.

Upon further investigation, Al-Khafafi found a significant difference in the ratios of simple repetition to complex repetition in the four longest chains in the two texts. In the four longest chains, the average ratio was 5:1 in the Arabic text, and only 1.4:1 in the English text. Moreover, in the Arabic text the average ratio in the four longest chains was twice as much as the average ratio for the texts as a whole, while in the English text it was less than the average ratio for all the chains (Al-Khafafi 2005:17). This finding suggests that writers of English texts, when using long chains of simple repetition, tend to mix these with complex repetition more than Arabic writers do. Al-Khafafi argues that this may contribute to the sense that Arabic texts rely more on simple repetition than English texts.

Al-Khafafi concedes that the corpus was fairly small (2005:20) and that the conclusions were tentative. His study nevertheless offers some explanation of how Arabic writers may use repetition as a rhetorical device to 'drive home their point' (2005:22). It also raises the question of how this may interfere in the English writing of Arabic speakers.

It is clear from these studies that, although repetition is an important part of lexical cohesion, if used inappropriately it can have a negative effect on the quality of a text. Indeed, as Witte and Faigley (1981:198) noted, frequent repetition does not improve the readability of a text. What is more important is how new information is introduced and ideas elaborated on.

This lends support to Hinkel's (2001) finding that it is not simply what cohesive devices are used but how they are used that is important. For instance, McGee (2009) argues that, while repetition may be a useful device, EFL writing often contains too much 'redundant repetition' (as opposed to repetition that serves to enhance the text). This can render the writing tedious and monotonous (McGee 2009:213–214). He suggests a reason for learners' reliance on repetition, namely, that in some cultural contexts (for example that of his Saudi university students) it may well be an accepted cohesive device. Another reason may be that students do not appreciate the importance of variety in writing. He also notes that insufficient attention is given to the use of superordinates. In most cases, a lexical chain usually moves from the specific to the general (Rover – Labrador – dog – animal). In some students' writing, however, this sequence may not be properly observed. For example:

One day, while I was out walking, I saw an animal. The dog looked lost, so I took it to the SPCA office. (My own example)

In this example, it is difficult to ascertain whether the dog in the second sentence is the same animal as that mentioned in the first sentence. Indeed, they appear to be unrelated sentences. However, if *dog* and *animal* are switched in the above sentences, it is perfectly clear that *animal* and *dog* are one and the same.

McGee's (2009) in-depth analysis of excerpts of his Saudi students' work focuses on two main types of lexical cohesion, namely, repetition and collocation. It also looks at how these devices were used in compositions by his students, who were at an intermediate level of English, and how one may address some of the problems associated with their usage.

McGee's analysis of these texts helps to highlight some of the problems associated with L2 writing at this level. Some of his findings, particularly those with regard to repetition, are discussed in §2.5.2.3 below.

One reason suggested by McGee (2009) for problems arising from the use of lexical cohesion is that students who derive most of their learning from textbooks and graded readers are not adequately exposed to authentic texts where there is a greater variety of lexical alternatives and options. This may well be true; exposure to authentic texts may benefit learners, particularly at intermediate and advanced levels. However, a brief perusal of the textbook which is the focus of the present study would show that not only are there a number of grammatical cohesive devices, but also some very useful lexical devices such as superordinates (*these beautiful trees – cherry blossoms*), and synonyms (*not able to – could not*). It is these devices, and the way in which they can improve one's writing, that, as Kolln (1999:98) points out, should be taught in writing classes, and which, together with more elaboration and the use of an appropriate register, may be a more important indicator of quality in English writing than the frequency of cohesive devices used.

The findings cited in this section are of particular importance to the present study, since they have addressed the question of inappropriate use of cohesive devices. Here again, however, it is difficult to determine whether these findings may be applied to learners who have only recently learned about compound and complex sentences and are grappling with applying these new structures.

2.4.2.3 Synonyms

Students often find synonyms particularly difficult, because there are relatively few words which mean exactly the same thing. McGee (2009) shows that the way in which synonyms are used is largely dependent on their context. There are, in fact, three basic factors to

consider. Firstly, there is direct versus indirect meaning (for example *lie* versus *misrepresent*); secondly there are attitudinal differences where the choice of a word also conveys the attitude of the speaker. For example, using the word *slim* to describe a person's size suggests that the speaker thinks the person is attractive and healthy, whereas the word *skinny* suggests that the person to whom the speaker is referring is possibly too thin. Thirdly, there are stylistic differences (for example, *the police* is formal, while *cops* is informal) (McGee 2009:215). Simply giving students lists of synonyms gives the impression that one word can easily be substituted for another. One therefore needs to show the contexts in which words with the same or similar meaning may be used. This is very useful, particularly with more advanced learners. However, the vocabulary of the learners in the present study was still quite limited, so it remained to be seen precisely how they would make use of synonyms.

All these studies clearly demonstrate that good writing is not dependent so much on an abundance of cohesive devices, as on their appropriate use. They suggest that simple rules such as 'Do not repeat the same words too much', or conversely 'Repeat key words and phrases', do not by themselves enhance the quality of writing. Rather, it is important to consider more closely precisely what kinds of devices students are using, how they are being used, and the context in which they are being used. Such information cannot be found simply by counting the frequency of cohesive devices. It is for this reason that a qualitative analysis of the texts will form an important part of the present study.

This section has examined studies of the different types of cohesive device used in EFL and ESL writing. It has also examined some of the problems found in the use of these devices. This leads to the question of the extent to which cohesion, and the choice of cohesive devices, is affected by the learner's first language. This is the focus of the next section.

2.4.3 Cohesion errors and first language (L1) transfer

A learner's first language has a profound effect on how he or she tackles the target language. This is no less true in his or her use of cohesive devices in writing. An important area of the study of cohesion concerns the way in which L1 conventions are transferred to the target language. This has to some degree been described with regard to the overall structure of texts (see §1.3.3 above), but is a problem at sentence and inter-sentence level.

Hinkel's (2001) study demonstrates the influence of the L1 on the use of cohesive devices. He found that the frequency of demonstrative pronouns in NNS writings exceeded those of NS writings. Demonstratives are probably seen as relatively simple to learn in English, and are also used in the languages (namely, Japanese, Korean, Indonesian and Arab) spoken by the participants of Hinkel's (2001) study. However, in his study, they were used somewhat differently in the students' mother tongue, and these usages were often transferred to English. The result was that it was often unclear precisely to what the demonstratives were referring:

It is not appropriate to criticize somebody's choice. Although my opinion is like that, this one point is to be stressed (Hinkel 2001:125).

For me, my major is piano performance I am not much satisfied with this (Hinkel 2001:125).

While some demonstratives appear to be used correctly (*this one point is to be stressed*), other demonstratives are not entirely clear (*Although my opinion is like that; I am not so much satisfied with this*). In the second excerpt in particular, it is not entirely clear to what *this* refers. Rather, as Hinkel (2001:125) notes, the demonstratives tend to refer more to general rather than to specific ideas or concepts.

In a more recent study, Liu and Qi (2010) examined how textual cohesion and coherence differed in the abstract writings of Chinese doctoral thesis writers and English-speaking academics. Abstracts written by 60 Chinese doctoral students were compared with those written by English and American researchers. Firstly, the use of cohesive devices as presented by Halliday and Hasan (1976) was recorded. Secondly, the use of repetition using Hoey's framework (1991) was investigated. Most interesting from the point of view of my study was the use of both Halliday and Hasan's (1976) framework and that of Hoey (1991).

Thirdly, a thematic progression (TP) model was used. This model divides each part of a sentence into a theme (the starting point) and a rheme (the main content of the sentence) (Li 2011:671). According to thematic progression theory, there are three main types of progression.

In the first type (known as the parallel TP pattern), the sentences have the same theme (T), but different rhemes (R) (Li 2011:672).

For example: He (T1) likes cheese (R1). He (T2 = T1) also enjoys exercise.

The second type is called the concentrated pattern, which has the same rheme but different themes (Li 2011:673). For example: I (T1) like the sea (R1). My family (T2) always goes to the sea (R2 = R1).

In the third type of TP, known as the continuous model, the rheme of one sentence becomes the theme of the next sentence, which has a new rheme. This new rheme becomes the theme of the following sentence (Li 2011:673):

I (T1) have always loved the sea (R1). It (T1 = R1) inspires me when I am feeling sad.

In the above example, *the sea* (R1) in the first sentence becomes the theme (*It*) in the next sentence.

Using the above framework, Liu and Qi (2010) compared the written English abstracts of Chinese L1 students to those of native English speakers as a means of comparing the frequency of cohesive devices and how they were used.

The results of Liu and Qi's study (2010) revealed a higher use of covert cohesion by native English speakers, which simply means that fewer conjunctions and more elision were used. Moreover, native English speaking writers used more direct repetition, complex repetition and paraphrase to continue ideas, whereas Chinese L1 writers used more direct repetition. What this suggests is that English NSs use a variety of ways to continue the topic; for example, in Liu and Qi's study (2010), the native speakers used a number of words to refer to a spacecraft, such as *space shuttle*, *spaceship*, *the shuttle*, *the ship* and so on.

Liu and Qi (2010) cited five main features with which the Chinese writers had difficulties, namely:

1. Overuse of cohesive devices
2. Too many overt links (such as overuse of conjunction and discourse markers)
3. Inappropriate repetition
4. NNSs tended to prefer concentration progression patterns and indirect statements, meaning that they changed the theme or topic of the sentence, but tended to keep the rheme or content part of the sentence.
5. Indirect or empty background information which was not relevant or appropriate to the conventions of abstract writing (Liu and Qi 2010:181), namely, *inter alia*, information on the writer's science interest and background, as well as 'general ideas for science development', and 'empty statements of important roles' (2010:179).

What Liu and Qi (2010) suggest is that many of these difficulties arise as a result of transfer from the L1. The Chinese writing style formed the basis of the Japanese writing style. As was noted in §1.3.3, Japanese and English writing styles differ markedly; Japanese structure is largely derived from the Chinese writing structure, which relies more heavily on indirect openings and statements. Moreover, as Liu and Qi (2010:181) note, English may be seen as a subject-based language, while Chinese is regarded as topic-based language. This means that, whereas English speakers place the subject of a sentence nearer the beginning of the text to give it prominence, Chinese speakers place the topic of the sentence, usually by means of an adverbial, nearer the beginning of the sentence. Thus, sentences often start with 'According to', 'Because of', 'Based on', etc. (Liu and Qi 2010:182). This can result in the subject of the sentence being lost, and the loss of a linear connection between one idea and the next.

As seen in Al-Khafafi's (2005) study in §2.4.2.2 above, another example of transfer from the native language lies in the use of repetition. Liu and Qi (2010:182) point out that in English, repetition is only acceptable when used as a rhetorical device; otherwise it is better to find a different way to repeat ideas. In Chinese, however, repetition is an important way of building coherence and cohesion. These patterns, too, are then transferred into Chinese speakers' English writing.

Liu and Qi's (2010) study is of great interest; firstly since two of its frameworks for analysis are the same as those used in this present study, and secondly because it also compares EFL students' writing with a native English model. However, given that the participants in the study were postgraduate students, it can be assumed that they had already been exposed to English for some years. Moreover, since the construction of a scientific abstract requires somewhat more than elementary-level proficiency, it can be safely deduced that the participants in this study were mature learners with relatively high language proficiency,

at least in the field of writing. All these points notwithstanding, it does raise the important question of how much transfer from the L1 takes place among high school learners.

Precisely which kinds of errors in cohesion are made at a particular level of learning was explored by Sadighi and Heydari (2012) in their analysis of cohesive errors made in writing by 67 male and female Iranian university students. Their aim was 'to empirically investigate, classify and analyze cohesive errors' (Sadighi and Heydari 2012:561). They wanted to establish, firstly, what the most frequently committed errors were; secondly, whether the differences that occurred could be attributed to the level of proficiency, and thirdly, whether the differences were due to L1 interference (Sadighi and Heydari 2012:561). Using the Oxford Placement Test to group the students according to their proficiency level into high-level and low-level groups, they had the students write a 200-word composition. A narrative topic was chosen because it was believed to be the easiest genre to write, thus allowing for the generation of more data.

Sadighi and Heydari's (2012) results showed that errors in reference were most frequently committed by low-level learners, followed by lexical errors and then conjunction. In the case of mid-level learners, reference errors were most frequent, followed by lexical and then conjunction errors. High-level learners made more lexical errors, followed by reference, conjunction and substitution (Sadighi and Heydari 2012:563–568). These results suggest that high-level learners were more exposed to the possibility of making errors, because they tended to use substitution more than did low-level learners. If higher level learners are familiar with substitution, they are more likely to make use of this device, thus increasing the frequency of substitution in their writing. Since there is a greater frequency of substitution in texts written by higher level learners, it follows that the likelihood of errors in substitution is increased.

An analysis of L1 interference in Sadighi and Heydari (2012) showed that the use of references was the most common error. This was because in Persian there is no distinction between pronouns, particularly personal and possessive pronouns, and demonstratives. This resulted in, among others, the following errors:

The old woman went to the hospital. He was sick (instead of the personal pronoun she) (Sadaghi and Heydari 2012:571).

As in the studies discussed above in §2.5.2, Sadighi and Heydari's (2012) study shows how learners at different levels of learning use different cohesive devices. This offers some support to Kolln's (1999) observations. Their study also demonstrates the importance of viewing errors in terms of L1 interference. The question that remains, however, is whether similar results may be found in the writings of Japanese junior high school learners. This question is more fully explored in Chapters 4 and 5.

2.4.4 Exophoric referencing

Another consequence of L1 transfer is the presence of exophoric referencing in ESL and EFL texts. As mentioned in §2.2 above, exophoric reference (namely, referring to something or someone outside the text) requires that both writer/speaker and reader/listener are aware of what is being referred to. Without this mutual awareness, communication breaks down. For this reason, some studies found exophoric reference to be a feature of unsuccessful writing. For example, in Olateju's (2006) examination of learner texts in a Nigerian school, one of the problems encountered was that of exophoric referencing (Olateju 2006:120). In certain circumstances, such as when both the reader and writer are aware of what is being referred to, this would be quite appropriate. However, if the reader has no idea of this presupposed referent, the writing can be very difficult to follow.

This use of exophoric referencing, as well as other problems in cohesion, Olateju (2006) believes, stems from interference from the L1. Olateju also notes a lack of coordination between different cohesive ties in his students' writing. This means that cohesive devices exist, but it is unclear precisely to what or whom the writer is referring. As Olateju notes, it is not only grammatical, but also lexical cohesive ties that are important (2006:127).

Olateju's (2006) study demonstrates the value of examining texts not only quantitatively but also qualitatively, so that one can determine not only what devices are present but also how these are used and how they have contributed to the success of the text. This observation lends support to the decision in this study to analyze both quantitatively and qualitatively. However, since Olateju's study took place in Nigeria, it is difficult to determine whether the exophoric references that he found were due to the peculiarities of the learners' particular L1 backgrounds, and whether similar results would be found among Japanese learners.

2.4.5 Focus on cohesion in Japanese high schools

It was noted at the beginning of this chapter (§2.1) that most studies of cohesion have been carried out with university students. However, since the main focus of this study is second-year junior high (equivalent to South African Grade 8) school learners, an examination of the studies that focus on high school learners is necessary.

Since assessment is an integral part of essay writing, it is hardly surprising that this should be the focus of study. Nakao's (2009) study focused on high school students (equivalent to South African Grades 10–12) and attempted to establish what language features assessors focus on when marking high school papers. She particularly wanted to know whether assessors focused on cohesion. Thirty-four texts were analyzed using Halliday and Hasan's criteria (1976). Four raters graded the texts, but with no criteria provided. According to Nakao's (2009) findings, great differences in the use of personal pronouns

were found (2009:53). Nakao suggests that this could have been due to the texts being personal anecdotes, a genre which tends to require or elicit a larger proportion of personal pronouns than would, say, an expository text. This is because a personal anecdote is most likely to refer to the first person (I), as well as to other characters in the story.

Nakao's study included an in-depth analysis of the highest and lowest rated texts based on Hoey's (1991) theory of the formation of links and bonds between sentences. An in-depth analysis of the strongest and weakest texts indicated a more consistent number of bonds between sentences in the highly rated essays, while the links tended to be more uneven and inconsistent in the lower rated essays. This suggests that, even though raters often cited other criteria (such as structure and organization) when accounting for the grades they gave, cohesion was clearly an important factor, although the extent to which the raters were actually aware of this was unclear. The study found that coherence was indeed an important criterion in the assessment of texts, and called for more consistent and constructive rubrics to be drawn up to help teachers in their assessment of high school learner compositions (Nakao 2009:57). According to Nakao's findings (2009), there were great differences between strong and weak essays in the use of personal pronouns (2009:53), but there was a similarity in the use of conjunctions and demonstratives.

Nakao's (2009) findings offer strong support for the study of cohesion in Japanese high schools. They also suggest that cohesion and coherence should be taught more explicitly in high school classrooms. She believes that exposing Hoey's (1991) theory to students could help them form a visual picture of the role that cohesion plays in 'connecting the meanings of sentences, resulting in coherence' (Nakao 2009:57). Nakao's study is most enlightening in highlighting the importance of cohesion in writing, as well as revealing some of the differences between more and less successful texts. However, the texts used in her study were still somewhat more advanced than those used in the present study. Moreover,

by not providing a model against which to compare the texts, it is not clear precisely how much the learner was expected to understand.

Okuda (2012) investigated the characteristics of lexical cohesion in the EFL writing of 30 second-year Japanese high school (South African Grade 11) EFL students' writing. The essays were rated by one Japanese rater and one English-speaking rater. The raters used a ten-point scale and were asked to specify the criteria used for rating the essays. The English-speaking rater tended to focus more on mechanical errors (grammar and spelling), and on ideas and their relevance to the topic. The Japanese rater focused on spelling, 'organization, content and general impression' (Okuda 2012:20–21). Focusing on five high-rated and five low-rated essays, Okuda's study used both quantitative and qualitative analysis. No significant differences were found between high and low-rated essays in terms of the frequency of lexical devices, except for synonymy. Again, the findings of the study suggested that it was not so much the quantity of cohesive devices as the way in which they were used which determined the success of a text.

Okuda found that repetition was the most used device in both the high-rated and the low-rated essays. The high-rated essays used more hyponymy, meronymy (that is, a word which denotes part of a whole, such as *finger* to *hand* [Hasan 1980:202]) and collocation. These essays also used a wider variety of lexical devices. The in-depth analysis of selected texts suggested a possible developmental sequence. Firstly, it highlighted the importance of developing subtopics around the assigned topic, then linking these topics and, finally, being able to elaborate on these ideas in detail. High-rated essays were able to do all three, namely, express an idea, connect ideas and elaborate. Writers of low-rated essays, according to Okuda (2012), were not able to use topic-related vocabulary to connect ideas. Essays that were rated in the middle range could connect ideas but not elaborate on them.

Okuda's study highlights some important issues that are relevant to my study. Firstly, there is the question of whether the frequency of cohesive devices reflects writing ability in any way. Secondly, it shows the importance of going beyond the statistics and examining more closely how the different devices are being used. This is what the qualitative section in my study set out to achieve. However, Okuda's study focuses on the writing of senior high school learners, rather than junior high school learners. Senior high school students already have three or four years of English learning behind them and, while Okuda's study addresses a number of important issues regarding Japanese high school EFL essays, there is still the need to investigate how students in the lower grades (J1–3), who have less vocabulary and fewer grammatical structures at their disposal, make use of cohesive devices in their written work.

2.5 SUMMARY AND CONCLUSION

This chapter began with a brief discussion on the difference between cohesion and coherence. It noted that one way of defining the difference was in seeing cohesion as the objective, quantifiable presence of links between sentences, whereas coherence was concerned more with the more subjective, qualitative sense of how the text as a whole is arranged and fits together.

The chapter then briefly provided some important developments in the study of cohesion, from Halliday and Hasan's (1976) taxonomy, through Hasan's (1980) revisions, to the work of Hoey (1991). It noted how Hoey's work offered a more objective, quantifiable means of resolving the cohesion–coherence dichotomy by focusing not on different types of cohesion as such, but on the way in which ideas are reiterated and form links between sentences.

The studies cited above have gone some way in addressing the questions posed by this study. Firstly, most of the studies have shown no significant difference between strong and

weak texts in the overall frequency of cohesive devices. This was seen in both comparisons between strong and weak NNS texts (Hinkel 2001; Castro 2004; Mojica 2006; Chen 2007; Nakao 2009) and comparisons between NNS texts and NS texts (Abadiano 1995). Regarding the kinds of cohesive devices found, these studies have shown that less successful writers tend to use more conjunctive devices (Abadiano 1995; Hinkel 2001; Liu and Qi 2010) and that writers of less successful texts tend to use more conjunctions for sentence transition than more advanced or NS writers.

The use of direct repetition was also found to be more frequent in weaker texts than in more successful texts (Castro 2004; Al Khafafi 2005; Mojica 2006; Chen 2007; McGee 2009). Liu and Qi's (2010) study also found more direct repetition in the writings of NNSs than of NSs. It was noted that repetition is not in itself an indication of weak writing, but rather the way in which it is used can make the difference between an informative text and a seemingly monotonous one.

McGee (2009) highlighted the use of synonyms, noting that their proper use may also be an important factor in the success of a text. Often learners are not aware of the different shades of meaning of synonyms. Difficulties with superordinates may also occur when writers fail to move from the specific to the general (for example, Labrador – dog – animal – creature).

The influence of the L1 was also noted as a factor in the writings of NNSs of English. This was noted by Hinkel (2001), who found that it was not always clear to what the demonstratives were referring. Liu and Qi's (2010) study found that the problems identified in the abstracts of Chinese postgraduate students' writing stemmed largely from conventions in Chinese writing. Moreover, Sadighi and Heydari's (2012) study of cohesive

errors in Iranian students' English writing also demonstrated the way in which L1 can influence the writing of English texts.

All of the studies used either Halliday and Hasan's (1976) taxonomy, or Hoey's (1991) repetition network, or both, as frameworks. This in itself makes them very valuable, since it is possible to see how these frameworks help to highlight some of the problems found in cohesion.

These and other studies go a long way to addressing the main questions of the present study. However, it has also been noted that some studies compared the writing of strong and weak NNSs, rather than comparing NNS writing to that of NSs. Secondly, many of the studies made use of native-speaking children (Abadiano 1995) or of university students (for example, Castro 2004; Mojica 2006; Chen 2007; Sadighi and Heydari 2012). This makes it difficult to apply their findings to the present study. Other studies used texts that were considerably more sophisticated than those found in Japanese junior high school classrooms. This can be seen in the studies of Al-Khafafi (2005), Liu and Qi (2010) and Sadighi and Heydari (2012).

The findings of the studies carried out in Japan (Nakao 2009; Okuda 2012) highlighted the need for cohesion to be taught in class and also supported previous studies. They also made use of both Halliday and Hasan's (1976) study, as well as that of Hoey (1991). This in itself was most useful in showing how these frameworks could be applied to learners in a Japanese school environment. However, as has already been noted, their subjects were Japanese high school students with considerably greater proficiency than the junior high school students in the present study. Moreover, both studies were more concerned with the differences between highly rated and low rated texts than with comparing the texts to a native ability model.

The present study, on the other hand, is more focused on what cohesive devices are found in the textbook and whether the learners are making use of these. It also aims to find out not simply which or how many devices are being used in the textbook and by the learners, but also how these are used to create a more effective text. Since my study uses the same frameworks as those cited above and was focused on a very different population, namely, younger elementary-level learners of English, it was believed that it would further deepen our understanding of how cohesion may contribute to the quality and coherence of a text.

The following chapter will discuss the approaches and methods used in the study, and how the data were collected, analyzed and interpreted.

Chapter 3: Method

3.1 INTRODUCTION

This chapter describes the methodology followed in this study. It discusses the focus of the research, the approach taken, the research framework, as well as how the study was carried out. It also describes the data collection and how the data were analyzed. In addition, it explains why these particular methods of analysis were chosen.

3.2 RESEARCH AIMS AND QUESTIONS

The aim of this study, as laid out in Chapter 1 (§1.5) was to investigate how the use of cohesive devices in the writing of second-year Japanese junior high school learners compared to that of the readings in their English textbook.

Main research question

- Assuming that the reading passages in an English textbook for second-year Japanese junior high school learners provide a model for the type and quality of writing that can be expected at this language level (namely, upper elementary), how does the use of cohesive devices in the textbook compare to that of the writing of learners?

The question is broken down into the following sub-questions:

1. What cohesive devices are found in the textbook readings (TBRs) and what cohesive devices are found in the student texts (STs)?
2. Is there a difference in frequency of use in the TBRs and the STs?
3. What patterns of reiteration occur in the TBRs?
4. How do the patterns of reiteration in the STs compare to those in the TBRs?
5. Which cohesive devices in the STs are used inappropriately or inaccurately?

6. What are the reasons for these errors?

3.3 RESEARCH DESIGN

The goal of the approach taken in the present study was to establish which devices could be used at the upper elementary level of language learning, whether these devices were in fact used and whether they were being used appropriately.

In the light of Seliger and Shohamy's (2003) division of research approaches into the analytical and the synthetic, the quantitative part of this study adopted a more analytical approach for reasons of feasibility. This means that the study was limited in its methods of data collection and the analysis of its data. While the quantitative analysis focused on the number and frequency of cohesive devices in both the student compositions and the textbook readings through the use of log-likelihood statistical tests, the qualitative analysis focused on the reiteration of words and ideas in the TBRs and STs, the patterns in which they occurred and what these revealed about the quality of the texts. The qualitative analysis also examined the errors in cohesion and offered reasons for these errors.

3.4 FOCUS OF THE STUDY

The main focus of the study was an analysis of the use of cohesive devices in the essays of Junior 2 (J2) students at a private school near Tokyo, Japan, and of readings from the textbook used by the students at the time the essays were written. The essays were written as part of an end-of-term test. The purpose of examining the textbook was to create a model against which the writing of the learners could be compared.

3.5 CONTEXT OF THE STUDY

The school at which the study took place is run by Jesuits, and is generally regarded as one of the top schools in Kanagawa, the second most populous prefecture in Japan after

the Tokyo Metropolitan area. In order to enter the school, students must write a rigorous set of entrance tests in Mathematics, Science, Japanese and Social Studies. Given their generally high levels of ability and motivation, there is often a good deal of original, creative thought and spirited participation among them.

English lessons start in the first year, and move at a much faster pace than in the public (i.e. government-run) schools. For example, by the end of the first year, students will have covered (that is, are expected to be able to recognize, spell and use in a sentence) about 1000 words and phrases, and will be familiar with present and past tenses. By the second year, they will be familiar with, among other structures, the present and past perfect tense forms and the passive forms, and will have covered about 2000 words of vocabulary. The chart below indicates the grammar structures that the learners in J2 had covered from the beginning of the academic year in April to the end of term test in July.

By the time these essays were written, the learners were familiar with the most basic grammar, including present and past tenses, as well as the structures found in Table 3.1. They had also had some experience writing simple self-introductory paragraphs, as well as summarizing a short fable in English. This meant that they were encouraged to use the most recently learned grammatical structures in their essays.

Table 3.1: Grammar structures covered in the first semester in J2 when the essays were written

Grammar/function	Example sentence(s)
Subject – verb – complement	The cherry blossoms look beautiful. I feel tired.
Subject + verb + direct object + complement	The book makes me happy. Keep your room clean.
Talking about the future using 'will'	I'll buy her some flowers.
Emphasizing	What a beautiful day! What a difficult question this is!
Subordinating conjunctions and clauses	Be careful while you're driving. Don't begin the test before the bell rings.
Type 1 conditionals	You won't get any dessert if you don't eat your vegetables.
Comparatives	The boy is bigger than the girl. The apple is more expensive than the orange.
Comparative questions	Who is taller, John or Peter?
Comparatives using as	Tom is almost as tall as his father. Kay can't run as fast as Maureen.
Superlatives	Which city is the biggest, Tokyo, Yokohama or Osaka?
More and more/-er and -er	He is getting taller and taller.
Tag questions	You're really excited, aren't you?
Relative pronoun: with	Mary often plays tennis with her friend Sally.
Relative clauses: subjective	The dog that lives next door was barking.
Relative clauses: objective	The man whom you see over there is a teacher.
Adverbs	Please speak more quietly.
Say/think/know that ...	Do you think that it will rain tomorrow?
Noun + to + verb	There are many things to see. Let's get something to drink.

The way in which the learners are taught should also be considered. Vocabulary is introduced in a number of different ways. I usually start by simply giving the class a wordlist

with the words in Japanese and English. After testing the learners on these words, I then do the reading section in which they appear. Other exercises that I have used include cloze exercises, crossword puzzles, as well as having the learners make up their own sentences using the new words, and then sharing these with each other. I also offer some strategies for remembering words, such as linking them to known words, using them in phrases and sentences, linking them to words in the mother tongue, or placing them in context. There is also a certain amount of grammar and translation, particularly in the English 1 class, which is almost always taught by a Japanese teacher. The quick pace of the lessons can sometimes mean that there is not as much time as one would like to consolidate work, or to prioritize the areas that need more focus. This was one of the motivations for carrying out this study: to find out more precisely the extent to which the writing of the learners could be compared to that of the textbook.

3.6 PARTICIPANTS

In this study, the population was defined as all students in the J2 level in the school at the time at which the essays were written (Brown 2004:111). This population was largely homogenous. Firstly, all the students were Japanese, and had been raised predominantly in the Japanese education system, which means that there were almost no ethnic or cultural differences between them. Secondly, they were all boys of the same age, eliminating any gender and age differences as a factor of extraneous variables. Thirdly, most of the students (with the exception of a few returnees from stays overseas) had been learning English for only about 18 months. However, even the returnees were generally more proficient in the BICS area of language learning (see §1.3.2) than in CALP skills. While many learners in my school were able to function in English on a conversational level, they tended to function cognitively in Japanese. Finally, the students were highly intelligent, academically focused and, by and large, highly motivated in their studies.

3.7 SAMPLING METHOD

If a sample is to reflect the population, 'each member of the population should have an equal chance of being selected for the sample' (Brown 2004:111). In the present study, the population from which the sample was drawn was an entire J2 grade level of 182 students, consisting of four classes of between 45 and 46 students each. After eliminating essays that did not reach the minimum of 60 words, and those students who chose not to participate in the study (see §3.9 for ethical considerations), a random sample of 36 essays, or nine per class of 45 or 46, was chosen by picking random student numbers. The 36 chosen essays were then randomly assigned new numbers to ensure privacy. A more detailed description of how the essays were written appears in §3.8 below.

Had the population been significantly heterogeneous in terms of ethnic or cultural background, gender or academic ability, or had the sample been much smaller with wide-ranging language abilities (Brown 2004:112), it would have been more appropriate to opt for a stratified random sample. However, since none of these conditions applied, and given the homogeneity of the group, I was confident that, with a one-in-five chance of selection, the sample was representative of the given population.

3.8 DATA

The two main sources of data were the student essays, or texts (STs), and the textbook readings (TBRs). These are described in detail below.

3.8.1 The student texts (STs)

In the second half of the trimester (from May to June), students had been given time for free reading. Most of the books read were graded readers. One week before the end-of-term test, the students were asked to choose one book they had read and to write a brief review of it. This review was then rewritten from memory in the final test. Prior to this,

they were provided with a model text, and one set of questions to fill in on the details of the book and another to help them with writing the essays (see Appendix 2). Questions included the name of the book, who the main characters were and what the learner's impression of the story was (see Appendix 2). This meant that, while they had practiced the task beforehand, the writing was done in a controlled setting, namely, a test room, and was, within the context of the test room at least, their own work. Learners were required to write a minimum of 60 words in the reviews. Essays which did not meet this minimum requirement were heavily penalized. There was no upper limit. I decided to choose 36 essays (nine from each class), since this represented one fifth of the population of 181 students. This sample was small enough so as not to be unwieldy, but large enough to reflect the population as a whole.

3.8.2 The textbook readings (TBRs)

Eleven TBRs were selected. These 11 texts were the first 11 texts in the J2 textbook, and comprised all the texts that had been read and studied by the learners in J2 up to the point of writing the essay test. (These reading texts are provided in full in Appendix 5.)

The textbook series used at the school is *Progress in English* by Robert M Flynn, SJ (2009). Flynn was a Jesuit Priest. Although his name still appears on the cover, the actual writing, editing and updating of the book is now carried out by an editorial committee of teachers from Jesuit schools in Japan, including the one at which this study took place. It is a series of five books, focusing on American English and culture in the first two books, and on British English and culture in the third and fourth books. Since it is compiled by teachers from Jesuit schools, there are some allusions to Christianity and a predominantly western ethos and content. However, there are also references to Japanese life, society and culture.

Each textbook is divided into 15 lessons. Each lesson is divided again into three sections called 'Scenes', each of which highlights one or two grammar points. Each lesson contains two reading passages, which at my school are taught by either one of the English teachers in the particular grade level. The English 1 teacher usually focuses on explaining the grammar and practicing the scenes, while the English 2 teacher usually follows a few steps behind, covering areas that could not be done by the E1 teacher, as well as practicing listening and speaking. The teachers usually divide the teaching of the reading sections between themselves, and it is up to them to decide who covers which areas.

The pacing of the series is, compared to many other textbooks in Japan, quite brisk. At my school, one textbook is covered per year, but many schools that use this series of books move through them at a much slower pace. For example, while students at my school have already covered Book 3 by the end of J3, many schools have only finished Book 2. The readings used in my study came from the first six chapters of Book 2, which are the lessons that had been covered by the time the learners' essays were written.

The textbook series is based largely on the audiolingual approach (which stresses the importance of forming habits through repetition [Richards and Schmidt 2002:39]), and thus includes many simple drill patterns. Each grammatical structure is contextualized in scenes that follow a simple story (such as a Japanese exchange student visiting the US), as well as stories, songs and poems. The reading sections introduce new vocabulary and use grammatical structures that have been covered previously. Each textbook is accompanied by a machine called a 'repeater', which the student keeps at home. The machine allows the student to listen to and repeat the material in the textbook, and to do listening exercises. A separate workbook for this purpose is also provided. Moreover, learners are given books written in Japanese which contain information on different aspects of customs and culture

in English-speaking countries, particularly Britain and the US. This helps them to build up background knowledge on the various topics discussed in the textbook.

In addition to grammar practice drills, the textbook contains passages of dialogue, as well as sections on reading, writing, listening and pronunciation. New words that appear in the main text are listed at the bottom the page on which they first appear, as well as in an index at the back of the book.

The reading sections are written in such a way as to reinforce the vocabulary and grammatical structures that have already been learned, and also to introduce some new vocabulary relevant to the topic under discussion. Since these passages were written purely with language learning in mind, they are not what some might call authentic texts. But they do give one a good indication of the kind of structures the students need to be familiar with at a particular school level. This means that learners are expected to be able to recognize and to use the structures accurately and appropriately, not only in grammar exercises, but also in compositions and in conversation.

The textbook is the main source of input and the primary focus of English study for the students at the school. For better or for worse, understanding and being able to apply the contents of the textbook is seen by many students as the main priority. The reading sections are therefore a useful reflection of what the students should be capable of at a particular point in the book. A focus on selected reading passages therefore would reveal the kinds of cohesive devices that learners could reasonably be expected to produce at any given point in the book.

3.9 PREPARATION OF DATA

The 36 essays, as well as the first 11 TBRs that had been studied by the students prior to the test were typed up and coded. The coding process is described in the next section.

For the qualitative component of the study, it was simply not feasible to analyze every text in detail. For this reason, two TBRs were selected. These were specifically chosen because one came at the beginning and one nearer the end of the semester, thus representing some growth in language proficiency. Moreover, neither of them contained dialog, which may have made them difficult to compare to the student texts.

As for the STs, six were finally selected for qualitative analysis. Two of these texts had been graded as high, and were deemed by me to be of a high quality for the level of the learners concerned. Two texts were graded average, and two were graded as below average. The criteria for determining the success of the texts were as follows:

Box 3. 1: Criteria for grading of student text essays

- Length: substantially more than 50 words
- Uses recently learned grammatical structures
- Good, clear telling of the story
- No incomprehensible or unclear sentences
- A lot of new information
- No or only a few simple grammatical errors
- No unnecessary repetition
- Contains a number of complex and/or compound sentences
- Elaborates on points made
- Contains important components of a review, namely:
 - Tells the story
 - Expresses an opinion or impression of the story
 - Explains why

Although I graded all the essays holistically, I kept in mind the criteria listed in Box 3.1. Grades were *excellent*, *very good*, *good*, *average*, *below average* or *poor*. It must be noted, though, that since many of the truly poor essays did not meet the required minimum of 60 words, and were not included in this study, most of the essays in this particular study were

either *excellent*, *very good*, *average* or *below average*. Having graded all 180 essays, and having selected a random sample of 36 essays, nine from each class of 45 learners, I then selected two which I had graded *excellent* or *very good*, two graded *average* and two graded *below average* for qualitative analysis. Apart from keeping in mind the grades, this was a random selection.

In the case of the six STs selected for qualitative analysis, in order to gauge the reliability of my rating, I asked four teachers at my school to grade the six texts holistically, based on their overall impressions. The raters were asked to rate the STs on a scale of 1 to 5, with 5 the highest score, and 1 the lowest. Since the goal here was simply to gain some consensus on the relative success of the texts, I kept the criteria deliberately vague to ensure that the scoring was based more on overall impression than on any specific criteria. The strength of the correlation between the four scores was then calculated using the correlation function on Excel. I simply calculated the average correlation between all five sets of scores. This correlation was 0.91, which indicated a strong consensus between the graders about what constituted a strong or a weak composition (The scores are shown on Table 4.16, §4.5.1 below).

3.10 FRAMEWORKS FOR THE STUDY

As noted in §1.7, two frameworks were used for the analysis in this study. The quantitative framework was based on the taxonomy of cohesion devised by Halliday and Hasan (1976), and the qualitative framework was based on Hoey's (1991) study of patterns of lexical repetition in texts. The discussion below examines the frameworks in detail, and the way they were adapted for the purposes of the study.

The first framework to be discussed is Halliday and Hasan's (1976) taxonomy of cohesive devices. Table 3.2 below was adapted from the original taxonomy, the full version of which appears in Appendix 1

3.10.1 Halliday and Hasan's taxonomy of cohesive devices (1976)

According to Halliday and Hasan (1976:4), cohesion refers to 'the relations of meaning within a text'. What this suggests is that one part of a text is dependent on another for its meaning. For example, one has to assume in a sentence such as 'They went to the beach' (my own example) that the pronoun 'they' refers to or presupposes some previously mentioned or acknowledged item; it cannot simply stand alone. This sets up a chain of words that refers back to previous sentences. The result is that sentences contain linkages of meaning between one another.

What follows is a description of the different cohesive devices as classified by Halliday and Hasan (1976). This description will cover first grammatical and then lexical devices.

3.10.1.1 Grammatical cohesive devices

The types of cohesion were briefly mentioned in §1.3.1. This section deals with grammatical cohesive devices, namely, reference, conjunction, substitution and ellipsis. The descriptions below also contain many important concepts that are crucial to understanding not only the present study but also previous studies on cohesion.

A Reference

There are two broad kinds of reference: exophoric and endophoric. An exophoric reference refers to something outside the text that has not been mentioned and needs to be constructed by the listener. This is not cohesive as such, but is an important concept to consider, since it often refers to the context in which the text is written. An example (my

own) of an exophoric reference would be of a son, having come home from a job interview, being greeted by his father with: 'So, how did it go?' While the interview has not been previously mentioned in this particular exchange, both father and son are clearly aware that 'it' refers to the interview, which exists outside the conversational text. The success of exophoric reference relies on the existence of shared information between reader and writer. If one party in the conversation or the reader of a text is unaware of the exophoric referent, communication breaks down.

Endophoric reference is concerned with the continuity of particular items within the text. For example:

Example 3.1

Mozart was born in 1756. He was the son of a musician. (My own example)

The 'he' in the second sentence can only refer to Mozart. Moreover, 'Mozart' is continued in the second sentence by the use of 'he'. As a result, the two sentences can be said to be tied together through the common reference to Mozart.

Endophoric reference can be either anaphoric or cataphoric. An anaphoric reference refers to something that occurred in any of the previous sentences. The 'he' in Example 3.1 above is an anaphoric reference, because it refers to *Mozart* in the previous sentence. A cataphoric chain, which is somewhat rarer, refers to something further ahead in the text. For example, the utterance, 'This is what he said' (my own example) is sometimes uttered or written before quoting a person.

According to Halliday and Hasan's taxonomy, reference is divided into three main types, namely, personal, demonstrative and comparative.

A1 Personal reference

Personal reference consists of two main subtypes, namely personal determiners (possessive adjectives such as *his*, *her*, *its* or the plural *their*), and personal pronouns (*he*, *she*, *it*, *they*). Personal pronouns can have exophoric reference, such as in a letter where the 'I' refers to the writer and the 'you' to the reader:

Example 3.2:

As many of you may know, I was married last month. My wife and I traveled to Mauritius for our honeymoon. She had always wanted to go there. (My own example)

In the above example, 'you' is addressing the reader of the text, while 'I' refers to the writer. Both of these appear outside of the text. The word 'she' on the other hand, refers to the writer's wife, who is mentioned in the text.

Reference may refer not simply to a particular word, as in Example 3.2, but also to an extended piece of text:

Example 3.3

A: Yesterday, I had a terrible day. I woke up late, was late for work, and then missed an important meeting. Finally the boss was angry and gave me a dressing down.

B: That does sound like an awful day.

(My own examples)

In the above extract, the 'that' in B's reply refers not simply to one particular word in the previous sentences but to A's utterance as a whole.

A2 Demonstrative reference

Demonstrative reference is, according to Halliday and Hasan (1976), a form of 'verbal pointing' (1976:57). Halliday and Hasan divide demonstratives into two broad types, namely, neutral, and selective. Neutral reference consists of the definite article *the*, which can in fact include both possessives (*his, her, my*) and demonstratives (*this, that, those*).

Selective demonstratives are, as the name suggests, more specific. They can either refer to near (*this, these*) or far objects (*that, those*). They are also subdivided into participant (*this, that, these, those*), and circumstance (*here, there, now, then*). This kind of demonstrative may be anaphoric or cataphoric determiners.

A3 Comparative referencing

Comparative referencing can be divided into general or particular references. General comparison reference can express identity (*same, equal, identical*), similarity (*such, similar*), and difference (*other, different, else*). Comparative referencing in particular is divided into numerative (*more, fewer, less*) and epithet (*better, so-so, more or less*).

B Conjunction

Halliday and Hasan (1976) describe conjunction in great detail, much of which is not relevant for the purposes of this study, simply because the learners' vocabulary was at a level where they had only recently learned the most basic conjunctions. Nevertheless, it is useful to look at the main categories, as well as how conjunction is defined by Halliday and Hasan (1976). Conjunction differs from reference, ellipsis and substitution in that, while these three usually refer back to particular ideas in the texts, conjunction indicates how 'what is to follow is connected to what has gone before' (Halliday and Hasan 1976:227), as in the example below:

Example 3.4

There was no answer. So I put the phone down. (My own example)

The *so* in the second sentence does not refer back to any particular part of the first sentence, or even the first sentence as a whole. It merely indicates that the second sentence is a consequence of the first.

Conjunction as a cohesive device is concerned with those conjunctions that refer to other sentences, rather than those that refer to a clause within a given sentence. This had to be borne in mind when collating the different cohesive devices in this study.

As has already been mentioned in §1.3.1.1, conjunction is divided into four main types, namely, additive (*and*), adversative (*but, however*), causal (*so, because*), and temporal (*then, later*). These are then further subdivided into a number of sub-sub categories. (For a full list of all conjunctive types, please see Appendix 1.) Since the students in this study had not, as yet, learned many of the words found in these subcategories, it is not necessary to go into any more detail here.

C Substitution and ellipsis

Halliday and Hasan (1976) devote some 140 pages to substitution and ellipsis. In some respects these two cohesive categories are fairly similar. Their main function is to save the speaker the trouble of having to repeat information which is already known to both parties (namely, speaker and listener).

Substitution uses a word in order to refer back to certain information, for example:

Example 3.5

- A: How do you know that John is going to meet us here?
B: He said so.

(My own example)

In the above sentence, the word *so* replaces the previous words *going to meet us here*.

Ellipsis, on the other hand, simply omits the unnecessary items:

Example 3.6

- A: How did you come to school this morning?
B: By bus

In the above dialog, Speaker B has omitted the words *I came to school this morning*, since they have already been uttered, and are thus known to both speaker and listener.

As far as this study is concerned, the learners were yet to learn substitution words, with the result that it was highly unlikely that many, if any, would be found. Moreover, while ellipsis is indeed an important device for avoiding unnecessary repetition in texts, it is largely found in dialog, such as the one quoted above. For these reasons, no more detail about these two devices was necessary for the purposes of this study.

3.10.1.2 Lexical devices

According to Halliday and Hasan, lexical cohesion is concerned with the 'selection of vocabulary' (1976:274). In this model, lexical cohesion is divided into two broad categories, namely, reiteration and collocation. Reiteration is then further divided into exact repetition, synonym, superordinate, and general words.

As was noted in §2.3 above, despite accounting for 40 to 50% of all devices, Halliday and Hasan (1976) devote only about 20 pages to lexical cohesion. This suggests that they saw grammatical cohesion as more important. However, the high proportion of lexical cohesive devices in Halliday and Hasan's taxonomy (Hoey 1991:9) highlights the importance of lexis in texts.

A General words

The first type of lexical cohesive device dealt with by Halliday and Hasan (1976) is general words. These are words that cover a wide range of categories. The most common examples are *creature* (non-human animate), *thing* (non-human inanimate), *stuff*, *business*, *move*, *place* and *idea* (Halliday and Hasan 1976:274). As with substitution above, these words had not as yet appeared in the textbook, meaning that most learners were not familiar with them.

B Repetition

Other types of lexical cohesion are concerned with reiteration. The first and most simple type is simple repetition:

Example 3.7

I found a wallet in the street. I then took the wallet to the police.

In the second sentence, the word *wallet* refers back to the wallet mentioned in the first sentence. The definition of repetition is expanded on later when discussing Hoey's framework in §3.5.1.2 below.

C Synonyms

Synonyms are words that mean more or less the same thing. One example of synonyms is *climb* and *ascend*. These may be used interchangeably, although how each one is used may be governed by outside factors such as the register of the text, whether it is formal or informal, or technical versus aimed at laypersons. Included in the synonym category are hyponyms, which are words that belong to a particular category. For example:

Example 3.8

Africa is famous for its many animals. Lion, leopard, rhino, water buffalo and elephant may be seen in many of the great national parks.

(My own example)

In the example above, the *lion*, *leopard*, *rhino*, *water buffalo* and *elephant* refer back to the *animals* in the first sentence. Since these are part of the greater category of *animals*, they can be said to have a hyponymic relationship, meaning that they are part of the category that went before.

Superordinates are words that form a more general class than the referent. For example:

Example 3.9

Yesterday, my Mini broke down. The car will take about a week to fix. This is difficult, since I am now without a vehicle with which to get to work.

(My own example)

In the above passage, *vehicle* is a superordinate of *car*, which in turn is a superordinate of *Mini*. Superordinates usually move from the specific to the general: Mini – car – vehicle.

D Collocation

Collocation refers to words which tend to occur together. For example, if the words *eat* and *apple* and *table* appear in the same text, it is most likely that the word *eat* will refer to the apple rather than to the table. However, the identification of collocation in a text is not as clear-cut as other cohesive devices. This is because deciding which words are actually collocations of one another is to some degree dependent on the reader's interpretation and knowledge of the subject of the text. Halliday and Hasan (1976) provide a number of different types of collocation which appear in Appendix 1. However, this detail is not necessary for this particular study.

One reason why this taxonomy was so useful in this study is that the researcher can be as general or specific as is necessary. If one simply wants to identify the main subcategories for conjunction, namely, additive, adversative, causal or temporal, this is quite possible. If a researcher wants more detailed information on conjunction, he or she may delve deeper, using further subcategories.

3.10.2 Adaptation and simplification of Halliday and Hasan's (1976) taxonomy

In the case of the present study, the upper elementary/low intermediate level of the textbook, and by extension of the learners, meant that many of the devices were not used in either the textbook readings or the student text parts of the taxonomy. For this reason, use of the taxonomy was somewhat simplified.

Table 3.2 shows a simplified version of the cohesion taxonomy set out by Halliday and Hasan (1976). Other sub-subcategories were collapsed into the sub-categories shown in Table 3.2. This was largely to ensure more accurate statistical analysis by having a larger number of items with which to work in each category. High-frequency devices in both the

STs and TBRs, or devices which warranted closer inspection, were broken down into sub-categories, whereas lower frequency devices were not broken down.

Table 3.2: Summary of cohesive types identified in this study

A: Grammatical cohesive types		
Cohesive type	Code	Example
1	Conjunction	
	additive simple	C11 <i>and, and also, nor, and ... not, or else</i>
	additive complex	C12 <i>add to that, alternatively, by the way</i>
	adversative	C21 <i>but</i>
	contrastive avowal	C22 <i>in (point of) fact, actually</i>
	contrastive	C23 <i>but, and</i>
	causal	C31 <i>so, then</i>
	temporal simple	C41 <i>then, next, before that</i>
	complex	C42 <i>soon, next time, next day</i>
	internal temporal	C43 <i>next, then, finally</i>
	here and now	C44 <i>up to now, at this point</i>
2	Ellipsis	E <i>I ate, but John didn't (eat)</i>
3	Reference	
	pronomial masculine	R11 <i>he, him, his</i>
	pronomial feminine	R12 <i>she, her, hers</i>
	pronomial neuter	R13 <i>it, its</i>
	pronomial plural	R14 <i>they, them, their, theirs</i>
	demonstrative near	R21 <i>this, these, here</i>
	demonstrative far	R22 <i>these, those</i>
	definite article	R23 <i>the</i>
4	Substitution	S <i>one, ones, so, do, be, have, do the same, likewise, do so, be so, not</i>
B: Lexical cohesive types		
Cohesive type	Code	Example
5	direct (simple) repetition	L1 <i>dog – dog</i>
6	synonym or near-synonym	L2 <i>big – large, animal – dog</i>
7	superordinate	L3 <i>dog – animal</i>
8	general words	L4 <i>creature, situation</i>
9	collocation	L5 <i>eat – food, dog – bark</i>

Source: Adapted from Halliday and Hasan (1976)

3.10.3 Coding of the data

In order to explain more clearly how the data were coded, a short four-sentence extract from one of the STs (Text 10: The Lost Key) is coded and analyzed below. The text excerpt appears in Box 3.2.

Box 3.2: Sample of Text 10 - *The Lost Key*

1. The name of the book is "The lost Key".
2. The story began in America, and the times is almost now.
3. The main characters are Biff and his friends.
4. The key which Biff had had magic power.

This coding was initially carried out by hand but was then captured as a full Microsoft Excel spreadsheet (see Appendix 4 for a sample of this coding). For the purposes of explanation, the coded section of the above extract is presented below:

Table 3.3: Sample of classification of cohesive types

Text 10: The Lost Key			
Sentence no	Cohesive item	Type	Presupposed item
1	the	R23.6	the book that I read (exophoric reference)
1	book	L1	book (exophoric reference)
2	the	R23.6	book
2	story	L3	book
4	Biff	L1	Biff

In sentence 1, two exophoric references are recorded. These are in fact mistakes, which were recorded separately and not counted when doing the statistical analysis, since there were no counterparts with which to compare them in the TBRs. They were, however, classified and analyzed in the qualitative analysis. In the column under Type, R23.6 indicates a definite article, L1 shows repetition, and L3 indicates a synonym. The presupposed items were also noted in the far right-hand column.

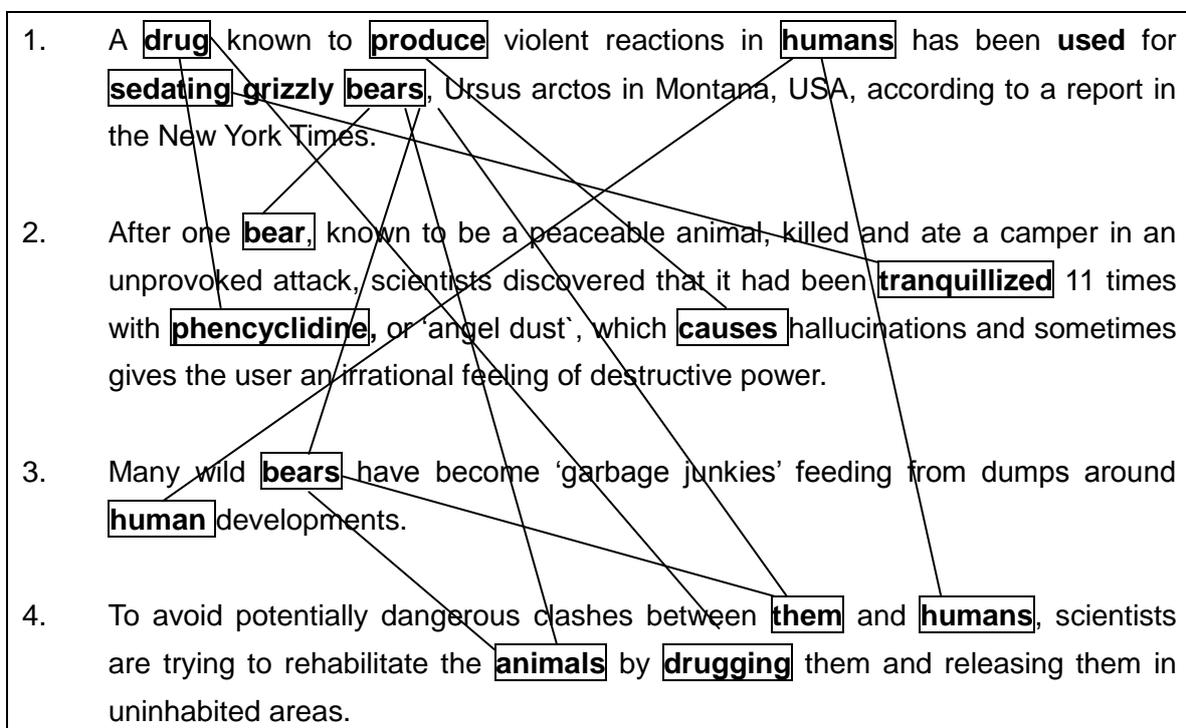
As can be seen, the sentence in which the tie occurs, the cohesive item and the code according to Halliday and Hasan's taxonomy (1976) are included.

Having described the main cohesive types in Halliday and Hasan's (1976) taxonomy, and how they were used to code the data, attention is now turned to the second framework in this study, namely, Hoey's (1991) model.

3.10.4 Hoey's (1991) patterns of lexis model

As noted in §2.3 above, Hoey's (1991) definition of repetition included not only direct repetition, but also, inter alia, synonyms, superordinates, antonyms and pronominals. With this in mind, a good starting point in a description of Hoey's work is to examine the following passage which appeared in his book:

Example 3.10



Source: Hoey (1991:37)

As can be seen from this example, words form links between sentences. It can also be seen that some sentences have more links than others. For example, there are four links between sentences 1 and 2, but only two links between sentences 1 and 3.

According to Hoey's criteria, three links (or repetitions) between any two sentences, no matter how remote, create a bond between these two sentences. Although sentences with two links may also be perfectly coherent, Hoey (1991) chose three links as the optimum number, because there were not too many or too few such bonds. In the case of two links, too many sentences would be classified as bonds, thus making it difficult to identify the key sentences in the texts. Using this model, Hoey (1991) demonstrated how sentences separated by thousands of words, or even in different volumes of text, could be coherent. To demonstrate this, he analyzed links from a three-volume work, *Masters of Political Thought* (Foster 1942 in Hoey 1991:79). Four links would probably have revealed too few examples for his purposes. Three or more links proved to be the number that provided both validity in the strength of the bonds themselves and reliability in terms of providing a reasonable number of such bonds. He showed how bonded sentences, even sentences that were separated by hundreds of other sentences, could be placed in juxtaposition to form a coherent text. Hoey's (1991) work further highlighted the presence of central sentences, which created a number of links and bonds, and peripheral sentences, which contained only one or two links at most. For example, in Example 3.10 above, sentence 1 forms a bond with sentences 2 and 4. These two sentences can be seen as the central sentences to the story. Sentence 1 does not bond with sentence 3. This means two things. Firstly, if one were to juxtapose sentences 1 and 3, there would be a sense that something was missing. Secondly, the information provided in sentence 3 is useful, but not critical to understanding the text. On the other hand, sentence 1 may be juxtaposed with sentence 4, and still make absolute sense.

Having examined the basic principles of Hoey's framework, we turn now to an examination of the different types of repetition described in his book.

3.10.4.1 Simple repetition

In simple repetition, a word is repeated with only minor grammatical alteration. For example, in Example 3.10 above, the repeated word (in this case, *bears*) is in fact referring to and includes one and the same *bear* (Hoey 1991:53). However, since this can only be achieved by lexical items, the repetition of grammatical items that refer to 'bear' is not counted.

3.10.4.2 Complex lexical repetition

Complex repetition occurs when two lexical items share a lexical morpheme, but are not formally identical. In Example 3.10 above, Hoey shows that while the words *drug* and *drugging* are morphemically similar, they are different in form (1991:55). Complex lexical repetition also occurs when the words are formally identical, but have different grammatical functions. In Example 3.10, *human* functions as either a noun (in sentence 1) or an adjective (in sentence 3).

3.10.4.3 Paraphrase

Paraphrase can be divided into simple and complex. Simple paraphrase occurs when one lexical item is substituted for another in context with a loss or gain in specificity, but with no discernible change in meaning (in Ex. 3.10 *drugging*, *sedating* and *tranquilizing* can be used interchangeably) (Hoey 1991:63).

Complex paraphrase occurs when two lexical items can be defined as including one another, even though they do not share the same lexical morpheme. In Hoey's model, complex paraphrase is restricted to three situations, namely:

1. *Antonymy*, meaning opposites such as *happy – unhappy* or *cold – hot*.
2. A *complex repetition* of another item, such as *writer – writings*, and then a simple paraphrase of a third: *writing – author*.
3. As can be seen, the presence of two types of links creates a third link, linking *writer* with *author* (Hoey 1991:64).

3.10.4.4 Superordinate, hyponymic, and co-reference repetition

If in a given text, *Mr. Smith* is first mentioned by name, and then is interchangeably referred to as either *the scientist* (more general) or the *biologist* (more specific), since it is clear that both are referring to *Mr. Smith*, it is possible to move back to a more specific reference after using a superordinate. When the repetition moves from the general to the specific (*scientist – Mr. Smith*), this is hyponymy (Hoey 1991:70).

3.10.4.5 Other types of repetition

Hoey treats personal pronouns as repetitions, so long as it is clear that they are referring to the same person or thing. However, first and second person pronouns (*I, you, we*) are only counted as repetition if they make it clear that they refer to someone within the text. If these pronouns refer to someone outside the text such as the writer, the reader or a third party, they are counted as exophoric references (Hoey 1991:71).

Demonstrative pronouns (*this, that, these, those*) are seen as entering into repetition links, but demonstrative modifiers (for example *this* in *this chair*) are not. For example,

Example 3.11

He bought four books. *These* he put into his bag. (My own example)

As can be seen, *these* in the second sentence refers to the books in the previous sentence, and would be counted as repetition, since it repeats the idea of the topic of *books* in the first sentence.

On the other hand, demonstrative pronouns are not included. This can be seen by looking at the following example:

Example 3.12

I have a number of favorite novels. *These books* have inspired me over the years.

(My own example)

In Hoey's model, the idea of *novels* is reiterated in the second sentence by the use of *books*. If *these*, which also refers to the *novels*, were counted, it would appear as if there were two reiterations where in fact there is only one. This is what Hoey refers to as the danger of 'double accounting' (1991:72).

In the case of substitution, links such as *one*, as in *the big one*, are counted as repetition (while including *the* in this instance would again constitute double accounting), as are the clausal *so*, as in *He said so*, as well as *another*, *the same* or *different* (Hoey 1991:74).

3.11 DATA ANALYSIS

As noted in §1.7, analysis was both quantitative and qualitative. The goal of the quantitative analysis was to establish precisely the types and frequency of cohesive devices to be found, firstly in the TBRs and secondly in the STs, and then to ascertain exactly how the presence or absence of cohesive devices in the essays resembled or differed from those in the textbook. Where necessary, I looked not only at the types and frequency of device, but also at how often specific words had been used in the STs and in the TBRs.

The qualitative analysis had three main objectives. The first objective, addressing Research Question 3, was to determine which actual devices were used in the cohesive chains and the patterns of reiteration in the model TBRs and, following from this, what the features of a successful text were, with particular regard to cohesion. As stated in Research Question 4 (§3.2), the second goal was to determine the extent to which six selected STs contained the same features as the TBRs. Addressing Research Questions 5 and 6 required examining in more detail the cohesive devices that were not used accurately in the STs, categorizing these errors, and offering possible reasons for students' difficulties with using certain devices.

3.11.1 Quantitative data analysis

The data were analyzed using the following procedures:

In addressing Research Questions 1 and 2, I used the following procedure:

- Raw scores of the coded data were tabulated so that it was possible to see precisely the number of devices present in each text (see Appendix 5).
- For the sake of simplicity, as well as to create larger samples for more robust statistical analysis, lexical items, which are not described in as much detail in Halliday and Hasan (1976), were coded L1–L5, ignoring the subcategories listed in the full taxonomy.
- The frequency of cohesive devices in the TBRs and the STs was compared using the log-likelihood statistical test (ucrel.lancs.ac.uk/llwizard.html). The reason for converting the numbers to frequencies rather than simply counting the raw scores was that the two corpora were of different sizes (that is, contained a different number of words). Converting the number of devices as a proportion of the number of words in each corpus meant that the proportions of each corpus could

be compared to one another. This test calculates the number of cohesive devices in the TBRs and STs as a percentage of the total number of words in each corpus. It then compares the true score (that is, the score that would theoretically be obtained if there were an infinite number of words) with the real scores, namely, the actual percentage of the total number of words made up by the cohesive devices to the total number of words. The likelihood score indicates whether the differences in frequency between the STs and TBRs are significant. A significant score would suggest that the differences between the two corpora are not due to chance (ucrel.lancs.ac.uk). One may then infer that the use of a particular device may be a factor in accounting for the success of the TBRs and the STs as a whole. A log-likelihood score does not account for the differences within the corpora. A comparison of the differences within selected texts was carried out in the qualitative analysis.

- The log-likelihood test was carried out using a program obtained from Rayson's LL calculator found on the Lancaster University (ucrel.lancs.ac.uk) website.

3.11.2 Qualitative data analysis

Before explaining the analysis in detail, it is necessary to explain an alternative way of referring to the words that were used in the texts, namely, referring to the number of words used in a given sentence as 'tokens' that could be counted. Secondly, any words that are repeated are referred to as 'types'. For example:

John lives in London. Although he has lived there for ten years, he still doesn't think of it as home.

The second sentence can be said to consist of 16 tokens. However, *he* appears twice, and refers to the same person, namely *John*. Thus one can say that the sentence contains 16 tokens, but 15 types (Nakao 2009:23).

The qualitative analysis was carried out as follows:

In addressing Research Question 4, the following procedure was followed:

- Two TBRs (TBRs B and J) were examined to determine the chains of reiteration, and the kinds of cohesive devices that occurred. Using the four lines in Text 10 (*The Lost Key*) (see Box 3.2 above), one of the cohesive chains would be as follows:

Sentence 1: The name of the book is “The Lost Key”.

Sentence 2: The story began in America, and the times is almost now

Sentence no.	Theme/idea	Cohesive type
1	book	
2	story	superordinate

Of course, in the above example there is only one instance of repetition visible, but when analyzing a full text, chains were often considerably longer.

- In order to determine what kind of patterns of reiteration could be found in the texts, I created grids similar the one shown below.

Diagram 3.1: Sample grid of reiteration patterns

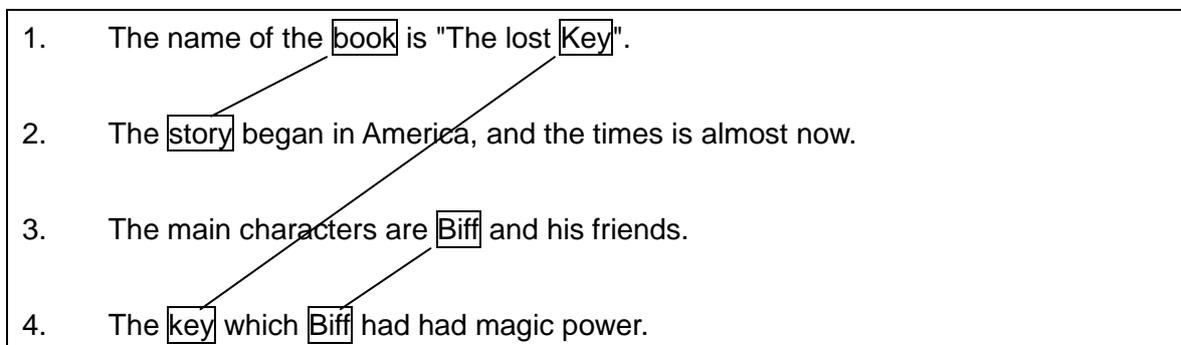
	1	2	3	4
book				
lost				
key				
America				
Biff				
friends				

The words in column 1 indicate the ideas or themes that were reiterated in the text. The row of numbers indicates the sentence number. Thus, it is possible to see in which sentence an idea or theme is reiterated. It is important to note that reiteration does not necessarily mean

repetition of the same word, but rather repetition of an idea. It may be any one of the types of repetition highlighted by Hoey (1991) and summarized in §3.10.4. Thus, reference, synonyms, superordinates and other types of reiteration were identified in the patterns of reiteration.

3. I examined the patterns of reiteration to determine how they were clustered, and where they appeared more spread out. I then investigated the meaning of these different distributions.
4. I also examined the links between sentences in order to see, firstly, if there were any bonds to be found (a bond being two sentences containing three or more links between them), as well as to determine which sentences contained words that were linked to a number of other sentences, and which contained links to only one or two sentences.

Diagram 3.2: The links between repeated words in Text 10 *The Lost Key*



As can be seen in Diagram 3.2, *book* is reiterated as *story* in sentence 2, *key* in sentence 1 is reiterated through direct repetition in sentence 4, and *Biff* in sentence 3 is repeated in sentence 4. This diagram helps us to see how the sentences are linked to one another, and which sentences contain the most links. It can also show how sentences with three or more links are bonded to one another.

5. Using the methods of analysis cited above, I was able to determine a number of common features in the TBRs that pointed to successful texts. These are listed in Chapter 4 (§4.4).

Having analyzed the TBRs, I then addressed Research Question 4 in the following steps:

1. I examined each of the six essays in turn to determine the extent to which they exhibited the six features identified in the TBRs. This was done by going through steps 1 to 5 described above.

Research Questions 5 and 6 concerned the errors in cohesion found in the STs. This was addressed as follows:

1. All the errors made by the students were noted during the coding procedure. They were then collected and assigned to one of four main categories, namely, exophoric reference, redundant repetition, unclear referencing, or conjunction.
2. The possible reasons for these errors were then discussed.

3.12 ETHICAL CONSIDERATIONS

The first ethical question which arose in this study was that of privacy. It was important that students' identity be protected (Hofstee 2006:118). The second question was one of consent. The third question concerned the extent to which participants were kept informed of developments in the study. Having had their essays analyzed and discussed in a thesis, they had the right to see the results for themselves.

The question of privacy was addressed first and foremost by assigning a unique, randomly selected code number to each essay used in the study. This number had no connection whatsoever to any student's name or seat number (a number assigned to every pupil each year according to the Hiragana syllabary).

Since the students were minors, I approached the school principal and asked him whether he felt that parental consent was necessary. I assured him that the name of the school, and the names and numbers of the students would not be revealed. He also asked me about the content of the essays, and whether these divulged any personal information. Once I had assured him that this was not the case, he deemed it unnecessary to consult the parents, and gave permission to use the essays.

As for the students themselves, I asked each one in the class to sign a letter of consent, explaining the basic aim and rationale of the study, how their essays were to be used, and assuring them of their anonymity (see Appendix 3). All students were asked to sign their names, and to then indicate with an O (Yes) or an X (No) – the traditional Japanese symbols of affirmation or negation – whether they agreed to have their essays used or not. Students were told a little about the study, and also assured that their names, student numbers and personal history would not be used. They were also told that they could, at any time, ask about the progress of the study, and that the results of the study would be available on request. This was done with the help of a Japanese teacher who reiterated my explanation to ensure that they completely understood what the study was about.

My one omission was in not explicitly letting the learners know that they were free to withdraw from the project at any time. However, some learners who had been selected later indicated that they did not wish to participate, and were immediately removed from the study and replaced with students who had given their consent.

3.13 SUMMARY AND CONCLUSION

This chapter focused on the methodology used to explore the use of cohesive devices and patterns of reiteration in textbook readings and student texts at a Japanese junior high school.

This chapter provided information on the population, the population sample, and how it was selected. It also explained how data were collected and analyzed. Finally, the ethical considerations were mentioned, as well as the measures taken to address these.

The following chapter will present the findings and discussion of these findings.

Chapter 4: Results and Discussion of the Findings

4.1 INTRODUCTION

This chapter provides the results of the analysis and a discussion of the findings. As mentioned in Chapters 1 and 3, the questions that it addresses are the following:

4.1.1 Reiteration of main research questions

Assuming that the reading passages in an English textbook for second-year Japanese junior high school learners provide a model for the type and quality of writing that may be expected at this language level (namely, upper elementary), how does the use of cohesive devices in the passages compare to that of the writing of its users?

The question is broken down into the following sub-questions :

1. What cohesive devices are found in the textbook readings (TBRs), and what cohesive devices are found in the student texts (STs)?
2. Is there a difference in frequency of use of TBRs and STs?
3. What patterns of reiteration occur in the TBRs?
4. How are the patterns of reiteration in the STs similar to or different from those in the TBRs?
5. Which cohesive devices in the STs are used inappropriately or inaccurately?
6. What are the reasons for these errors?

Questions 1 and 2 are quantitative, while questions 3 to 6 are qualitative in nature.

4.2 FINDINGS RELATING TO QUESTION 1

This section relates to Question 1: What cohesive devices are found in the TBRs and what cohesive devices are found in the STs?

Table 4.1 shows the number of cohesive devices identified in the STs and TBRs. At this point, all the TBRs and all the STs had been combined to form two corpora of text. This table simply shows the number of cohesive devices in each corpus. The percentage column indicates the frequency of devices as a proportion of the total number of words in each corpus. The figures as presented below are simply the raw numbers as they were identified and counted.

Table 4.1: Number of cohesive devices identified in each corpus

Type	TBRs		STs		
	Count	Corpus %	Type	Count	Corpus %
Conjunction	19	0.67	Conjunction	56	1.69
Ellipsis	38	1.35	Ellipsis	13	0.392
Reference	173	6.14	Reference	266	8.02
Substitution	6	0.21	Substitution	2	0.06
Subtotal	236	8.38	Subtotal	337	10.16
Repetition	267	9.48	Repetition	318	9.59
Synonyms	33	1.17	Synonyms	10	0.30
Superordinates	42	1.49	Superordinates	45	1.36
General Items	0	0.00	General Items	1	0.03
Collocation	143	5.08	Collocation	152	4.58
Subtotal	485	17.22	Subtotal	526	15.86
Total	721	25.59	863	863	26.03

Note: TBR corpus comprised 2817 words and the ST corpus 3316 words.

In both the TBRs and the STs the three most frequently occurring cohesive devices are repetition, reference and collocation. These three devices alone account for 81% of all the devices in the TBRs, and 85.2% of all devices found in the STs. The results as they appear above were somewhat unexpected. Given that the textbooks were written by native, or at the very least, near-native speakers of English, while the STs were written by learners who had, by and large, only been studying English for a year and a half, one would have

expected a far greater discrepancy in the types of devices and the frequency of the different cohesive devices used. Yet the two groups of texts appear remarkably similar in both distribution and types of device. This suggests that, at face value at least, the number and type of cohesive devices is not necessarily a factor in the quality of a text. The results also indicated that the learners were to some extent internalizing and using the structures as taught in the textbook in their writing.

The fact that these scores represent each corpus as a one body of text raises three issues. Firstly, they do not establish whether any differences that do exist between the two bodies are in fact significant. Secondly, the numbers do not give any indication of the actual words that were used. Thirdly, while there appear to be strong similarities between the STs and TBRs in the frequency of different cohesive types, the figures do not reveal how these devices are used.

The most noticeable feature of Table 4.1 is the predominance of lexical devices, which constitutes 17.22% of all the words in the TBRs, and 15.86% of the words in the STs. Furthermore, the preponderance of repetition in both types of text is not surprising, given that the writers of this textbook are limited to the vocabulary previously covered in the series, and a few new words. They are also trying to repeat words and expressions in order to review and reinforce them.

In order to address the first of these issues, namely, whether there was any significant difference between the groups, a log-likelihood analysis was carried out (as described in §3.11.1). A significant difference in the frequency of cohesive devices between the TBRs and the STs would suggest that these differences were most probably not due to chance alone. It could then be deduced that the frequency of a given cohesive device would be a factor in the quality of the text.

The second issue of precisely which words were used in each body of text was then examined in the light of the statistical analysis. These findings are presented in the next section (§4.3). The issue of how the devices were used is addressed in §4.5 and §4.6 below.

The following section examines the significance of the differences between numbers of CDs in the TBRs and in the STs. Grammatical devices will be discussed first, followed by lexical devices.

4.3 FINDINGS RELATING TO RESEARCH SUB QUESTION 2

This section discusses the findings relating to research sub-question 2: Is there a difference in frequency of use of CDs in TBRs and STs?

The most important feature of these results is the similarity between the ST and the TBR results. This can be seen in Table 4.2 below:

Table 4.2: Log-likelihood results for all devices

	TBRs	Corpus %	STs	Corpus %	Log-likelihood
Total CDs	721	25.59	863	26.03-	0.11
Total words	2817		3316		

A log-likelihood calculation showed that there was no significant difference between the TBRs and the STs in the frequency for the total number of devices (LL = 0.11, $p < 0.05$). The - in the ST percentage column indicates that there were marginally fewer cohesive devices used in the STs than in the TBRs. Since the log-likelihood is only 0.11, which is not close to the critical value of 3.84, the difference cannot be said to be significant. However, further investigation into individual devices revealed a more complex picture.

4.3.1 Grammatical devices

The results of the analysis of the difference in frequency of the grammatical devices are reflected in Table 4.3 below.

Table 4.3: Log-likelihood results for grammatical devices

	TBRs	Corpus %	STs	Corpus %	Log-likelihood
Conjunction	19	0.67	56	1.69	***13.54
Ellipsis	38	1.35	13	0.39+	****17.22
Reference	173	6.14	266	8.02-	**7.6
Substitution	6	0.21	2	0.06+	2.8
Subtotal	236	8.38	337	10.13-	*5.23
Total words	2817		3316		

95th percentile; 5% level; $p < 0.05$; critical value = 3.84*

99th percentile; 1% level; $p < 0.01$; critical value = 6.63 **

99.9th percentile; 0.1% level; $p < 0.001$; critical value = 10.83 ***

99.99th percentile; 0.01% level; $p < 0.0001$; critical value = 15.13 ****

There is a significant difference between the TBRs and the STs in the frequency of conjunction (LL = 13.54, $p < 0.001$), ellipsis (LL = 17.22, $p < 0.0001$) and reference (LL = 7.6, $p < 0.01$). STs featured relatively more examples of reference and conjunction devices, but fewer ellipses relative to the TBRs. As can be seen, the STs used significantly more grammatical devices in total relative to the TBRs (LL = 5.23, $p < 0.05$). Since grammatical devices are some of the first words learned and acquired, it stands to reason that learners would rely on these for the creation of cohesion.

The significantly higher use of ellipsis in the TBRs can be explained by the fact that these texts contained a number of dialogs (as opposed to prose), which by their nature make greater use of ellipsis. On the other hand, the STs (being book reviews which were simply relating a story and then offering an impression) contained very little dialog. In this study,

this factor needs to be considered when assessing the importance of ellipsis in the quality of the texts.

The use of conjunctions in the STs and TBRs is discussed in more detail in § 4.3.1.1 below.

4.3.1.1 Conjunctions

Conjunctions presented something of a problem in that many conjunctions were used erroneously in the STs. Eventually, I decided that the use of conjunction in the STs, whether erroneous or not, needed to be compared to the use of conjunction in the TBRs. A fuller discussion of errors in conjunction use can be found in § 4.6.3.

Although the TBRs used fewer conjunctions relative to STs (see § 4.3.1), they used a greater variety. On the other hand, the use of conjunctions in the STs was limited to a few basic types. In order to ascertain more clearly how the use of conjunctions may have influenced the quality of the ST texts, I investigated the actual words used in both the TBRs and the STs. The results are presented below:

Excluding the three conjunctive errors mentioned above, the STs used nine different conjunctions or conjunctive phrases in total: *after that; after; just at that time; but; and; then; so; finally; also*. However, of all the conjunctives identified in the STs (correct and erroneous), *but* was used 12 times, *and* nine times, *then* eight times and *so* six times. This means that 34 of the 44 conjunctive tokens were made up of only four types.

The TBRs used conjunctives far more sparingly (only 19 times overall in 2817 words) (LL = 10.43, $p < 0.001$). They also used a greater variety (14 different types) spread more evenly throughout the texts, meaning that there were eight conjunctions used in the TBRs that did not occur in the STs. Moreover, only six conjunctives occurred in both the STs and the TBRs. The log-likelihood calculation shows a significantly higher use of the words *and, but*

and *then* in the STs to start sentences and to indicate time.. In the STs, the word *and* was used four times to mean *then*.

On the other hand, the TBRs used nine conjunctions (*and so, also, of course, now, not only ... but also, next time, later, in fact*) that did not appear in the STs. There are two possible reasons why these were not used in the STs. The first is that the nature of the assignment did not allow for the use of these conjunctions. The second is that these conjunctions were not yet a part of the learners' active vocabulary.

4.3.1.2 Reference

The results indicate that the STs over-used demonstratives when compared to the TBRs (LL = 12.14, $p < .001$). However, none of the scores for the individual subtypes of demonstratives were significantly different. Thus, the significant difference between the STs and the TBRs with regard to demonstratives as a whole is probably due to the cumulative difference between the two groups. The question then arises of why there was a higher frequency of demonstratives than in the TBRs. In order to account for this difference, it is helpful to look more closely at the words to which they referred and their roles in the text.

Table 4.4 below shows the demonstratives in the STs, and the actual words to which they referred.

Table 4.4: Words and phrases referred to by the demonstratives in the STs

Demonstrative	Referent	Frequency
that	button	1
this	the rocket	1
this	book	19
this	donkey	1
this	in prison	1
this	of the broken shower	1
this	story	18
it (should be <i>that</i>)	the children getting sadder	1
that	they cleaned the park	1
that	the lift	1
these	the old toys	1
this	the title of the story	15
Total		61

While the differences between the STs and the TBRs in the use of demonstratives were statistically significant, the numbers do not tell the full story. Table 4.4 shows that, of the 61 near and far demonstratives identified, 19 referred to the word *book*, 18 to the word *story*, and 15 to the title of the book. This means that 85% of the demonstratives used referred to only one of three different words. Other than these three referents, there were only nine other referents. This shows that in most cases the use of demonstratives was limited to repeating parts of the model book review (see Appendix 2) that had been distributed before the test. Without the reference to *a book*, *a story*, or the story's title, there are in fact relatively few independent attempts at making use of demonstratives.

On the other hand, as Table 4.5 below shows, the demonstratives in the TBRs referred to 48 different types, phrases or sentences in only 11 texts. Only six words were referred to more than once using demonstratives and, of these, one was referred to three times and one four times. This suggests that, in more successful texts, demonstratives refer back to a wide variety of previous words and ideas.

Table 4.5: Words and phrases referred to by demonstratives in the TBRs

Frequency	Word	Text
1	Potomac River	B
1	Promote friendship	
1	River	
1	Manhattan	C
1	New York	
1	coast	E
1	explorers	
1	group of people	
1	states	
1	enemies	
1	Pocahontas	
1	Britain	F
1	goods	
1	a cup of tea	
1	threw tea in harbor	
1	ships	
1	Just put John Hancock here	G
1	cheeseburger	
1	cola float	
1	hot dog	
1	one Saturday	
1	food	
1	20-dollar bill	I
1	a picture of the White House	
1	he became President in N.York	
1	pamphlet	
1	White House	
1	Merchant	J
1	Slavery	
1	when A. Lincoln was young	
1	woman	
1	"I have a dream"	K
1	bus	
1	lack of rights	
1	silent protest	
1	Sentence 15	E
1	Sentence 17	H
1	Sentence 21	I
1	Sentence 22	K
2	colonies	D
2	colonists	E
2	Capt. John Smith	
2	chief	
2	Congress	G
2	Christ and Gandhi	K
3	Colony/ies	D; E
4	trees	B
6	the New World	D; E
64		

The log-likelihood was calculated for the number of different types referred to using demonstratives; 47 in the TBRs and 12 in the STs. The log-likelihood ratio was + 28.29, indicating a significant overuse in the TBRs relative to the STs. When looking at the

differences between the TBRs and the STs in how the demonstratives were used, it is clear that there existed a large gap between the textbook and the learners in terms of the variety of types to which the demonstratives referred.

Table 4.6 below identifies the definite article referents that occurred two or more times.

Table 4.6: Definite article referents that occurred two or more times

STs	Frequency	TBRs	Frequency
children	10	throw the tea into the harbor	3
book	8	The Declaration of Independence	2
boy	7	of Congress	2
rocket	5	chief	2
story	4	colony	2
key	3	colonists - group	2
money	2	cola float	2
leopard	2	coast	2

Of the 68 definite articles identified in the STs, 34 referred to one of only five types: *children* (10), *book* (8), *boy* (7), *rocket* (5) and *story* (4). One possible reason for the relatively high frequency of words such as *children*, *book*, and *boy* is that many of the books that learners reviewed belonged to the same series of graded reader. These graded readers contained many of the same characters, namely, a group of children that had all kinds of adventures.

In the TBRs, on the other hand, the 11 texts contained 31 different referents. Although the phrase *the Boston Tea Party* was referred to three times with the definite article, this was the highest frequency for any one particular token or idea. Six other types were referred to twice (*Declaration*, *chief*, *colony*, *colonists*, *cola float*, *coast*). This suggests that the definite article was used not only more sparingly in the TBRs, but was also used to refer to a wider range of types than was the case in the STs.

4.3.2 Lexical cohesive devices

The log-likelihood tests showed slight overuse by the TBRs, but the difference was not significant (LL = 2.06, $p < 0.05$) with no significant difference between the TBRs and the STs in the total frequency of lexical devices. While this seems to suggest that the frequency of lexical devices was not a factor in the creation of a successful text, a breakdown of the individual lexical devices, as well as an examination of precisely what actual words were used revealed a more complex picture.

Table 4.7: Log-likelihood results for lexical devices

	TBRs	Corpus %		STs	Corpus %	Log-likelihood
Repetition	267	9.48	Repetition	318	9.59+	0.02
Synonyms	33	1.17	Synonyms	10	0.30+	17.01****
Superordinates	42	1.49	Superordinates	45	1.36	0.19
Collocation	143	0.00	Collocation	152	0.03	0.77
General Items	0	0.00	General Items	1	0.03	0.04
Subtotal	489	17.36	Subtotal	526	15.86+	2.06
Total words	2817			3316		

99.99th percentile; 0.01% level; $p < 0.0001$; critical value = 15.13 ****

The log-likelihood ratios showed no significant differences in the use of repetition, superordinates, general words, and collocation. (Since general words had not yet appeared in the textbook, I did not examine these further.)

There was a significantly higher frequency of synonyms in the TBRs than in the STs (LL = 17.01, $p < 0.0001$). This suggests that more successful texts make more use of synonyms.

4.3.2.1 Repetition

As mentioned above, log-likelihood results revealed no significant difference between the TBRs and the STs in the use of repetition. In this study, this suggests that the learners may be mirroring the TBRs in this regard. However, in order to gain a more precise idea of how

direct repetition was being used, I examined exactly which parts of speech were being repeated. This deeper analysis of the words themselves revealed a more complex picture.

Table 4.8 below is a breakdown of the frequency of parts of speech that were repeated in both the STs and the TBRs.

Table 4.8: Repeated words according to parts of speech log-likelihood table

	TBRs	Corpus %	STs	Corpus %	Log-likelihood
Verbs	29	1.03	33	1.00	0.02
Proper nouns	85	3.02	81	2.44+	1.85
Common nouns	104	3.69	203	6.12-	18.40****
Abstract nouns	13	0.46	0	0.00+	20.23****
Stock phrases¹	8	0.28	1	0.03+	7.40
Adverbs	3	0.11	1	0.03+	1.40
Adjectives	25	0.89	7	0.21+	13.89***
Subtotal	267	9.48	326	9.59-	0.02
Total words	2817		3316		

99.9th percentile; 0.1% level; $p < 0.001$; critical value = 10.83 ***

99.99th percentile; 0.01% level; $p < 0.0001$; critical value = 15.13 ****

a (+) indicates overuse in the TBRs relative to the STs, while a (-) indicates underuse.

The first noticeable feature of Table 4.8 is the strongly significant difference in the use of common nouns (LL = 18.4, $p < 0.0001$). In fact, the log-likelihood score indicates a certainty of 99.9 that that the significantly greater repetitions of common nouns in the STs suggest that the TBRs repeat a wider variety of words.

The frequency of abstract nouns, stock phrases and adjectives in the TBRs were fairly low, but this is most likely due to the fact that the textbooks were at an elementary level and were only gradually increasing these parts of speech as the book progressed. However, in

¹ A stock phrase is an expression which can be remembered as if it were a whole word, for example, 'How are you?', 'What's the matter?', or in the case above, 'Let's go.'

the TBRs there was still a significantly higher frequency of repeated abstract nouns, stock phrases and adjectives. This suggests that the TBRs made use of a wider variety of parts of speech, while the STs were, by and large, more dependent on the repetition of common nouns.

This spread of repetition over different parts of speech means that, while common nouns may be the most prevalent part of speech used in repetition, they were not the dominant part of speech in the TBRs. Reiteration of ideas occurred across a spectrum of parts of speech rather than simply in the repetition of common nouns. Another possible reason for the spread of repetition over various parts of speech could be that the writers of the TBRs consciously repeated words as a means of reinforcing their acquisition. The use of one cohesive device (in this case repetition) using a variety of parts of speech also makes the text more interesting and more readable, while reiterating and reinforcing important ideas.

The repetition of abstract nouns warrants further attention. Table 4.8 above shows significant overuse of abstract nouns in the TBRs compared to the STs. These nouns, although by no means the most prevalent, perform an important role by not simply describing concrete situations but in explaining and reinforcing ideas. However, the learners either chose not to or were unable to use abstract nouns in their writing, despite having encountered them in the TBRs they had studied.

4.3.2.2 Synonyms

The significantly higher frequency of synonyms in the TBRs was to be expected, because the writers of the textbooks would have had a greater awareness of the synonyms available for use, while for most learners many synonyms would still have been part of their passive vocabulary. Although synonyms represent a small proportion of the cohesive devices used, their use is important in adding variety to a text. As for the STs, the use of 10 synonyms in a

corpus of 3316 words indicates that the learners as a group had not, as yet, mastered the use of synonyms. The possible reasons for this are dealt with in the following chapter (§5.2.2.2).

4.3.2.3 Superordinates

One of the more surprising results was the lack of a significant difference in the use of superordinates. Given that the textbook writers had a better grasp of English and the options available to them, one would have expected there to be a higher frequency of superordinates in the TBRs. For this reason, I decided to examine more closely precisely which superordinates were used.

I looked more closely at precisely which words were used in both the STs and the TBRs. The number of tokens, and the number of different types, found is presented in Table 4.9 below.

Table 4.9: Type-to-token ratio for superordinates used in the STs and TBRs

	Types	Tokens
STs	22	44
TBRs	36	42

What Table 4.9 shows is that the learners were limited to using superordinates for particular words only. This is confirmed in Table 4.10 below where the superordinates used more than once in the STs and in the TBRs are listed. Although I had in my lessons drawn the learners' attention to superordinates as a useful way of referring to what had come before in a text, they had not had much practice in using them in their writing, or in exploring the extent to which they could be used. This is borne out in the results below.

Table 4.10: Superordinates used more than once in TBRs and STs

TBRs		STs	
Superordinates	Occurrences	Superordinates	Occurrences
Man / men	4	story	11
people	2	book	7
New World	2	work	2
		three	2
		runs away	2
		children	2
Total	8	Total	26

As can be seen in Table 4.10 above, 18 occurrences of the 44 superordinates found in the STs comprised only two types, namely, *story* and *book*. In most cases the ST superordinates referred to the titles of the stories being reviewed. Another four superordinates (*work*, *three*, *runs away*, *children*) were used twice. In the TBRs, the superordinate used most was *man/men* (four times). Apart from this example, only two other words were used twice (*people* and *New World*). This indicates that the TBRs made use of a far greater variety of types than the STs.

The overuse of superordinates in the TBRs, though not significant, suggests that, at face value, the frequency of superordinates is not a factor in determining the quality of a text. However, when the most frequently occurring superordinates in each of the corpora were compared, 11 cases in the STs involved only one word (*story*), whereas the most frequently occurring superordinate (*man/men*) in the TBRs occurred only four times. Moreover, the STs contained 26 superordinates that occurred more than once, of which 18 were one of only two words (*book* and *story*). In the TBRs, only eight superordinates occurred more than once.

The results above clearly point to the learners' limited vocabulary at this stage. While it may be possible that they understood words, they were at this point unable to make use of the words that had appeared in the book. This discrepancy between the two corpora in the number of different types used indicates one area where the learners were lacking in sophistication and practice. It is clear that they needed to be made more aware of the way in which superordinates can be used to add variety to the texts. (For a full list of superordinates found in the TBRs, see Appendix 7.)

4.3.2.4 Collocation

While more collocation occurred in TBRs than in STs, the difference was not significant (0.77). This means that, in this particular study, the frequency of collocation as a significant factor in determining the quality of the texts could not be conclusively established. However, as was mentioned in Chapter 3 (§3.10.1.2D), collocation is a controversial cohesive device and is dependent to some extent on the subjectivity of the scorer. For this reason, the results for collocation need to be treated with caution.

4.3.3 Summary of quantitative results

In the TBR corpus there was significant overuse of ellipsis and synonyms, while in the STs there was a significant overuse of conjunction and reference. However, these results did not provide the full picture. While the statistical analysis indicated no significant difference between the two bodies of text in the use of cohesive devices in general, the analysis of the actual words used reveals a significant gap between the TBRs and the STs in the type of cohesive devices used. This difference was highlighted further in the qualitative analysis (see § 4.4 and §4.5).

The next step in the analysis of the data was to look more closely at how the cohesive devices in the two corpora were used. In order to do this, the definition of repetition is expanded to include repetition of not simply the same words but also of the same ideas.

In § 4.4 and §4.5, which form the qualitative part of the study, a second framework based on the work of Hoey (1991) is used to show not only which cohesive devices were used in the corpora but also how these devices combined to enhance a text's cohesion. The extent to which the learners' texts compared to the TBRs in the formation of patterns of reiteration is also discussed.

4.4 FINDINGS RELATING TO QUESTION 3

This section includes a discussion on the findings related to Question 3: What patterns of reiteration occur in the TBRs? Although three devices, namely, repetition, reference and collocation together accounted for most cohesive devices in both the TBRs (82%) and the STs (87%), the quantitative analysis did not reveal the extent to which the relative position of these devices in the texts themselves contributed to the success (or otherwise) of the text. Nor did it reveal how the devices created the links and chains of meaning that help to make a text both cohesive and coherent. This is the focus of this qualitative part of the study.

This section focuses on how ideas were reiterated using the devices discussed in the previous section. With regard to the TBRs, the analysis aimed to establish how the use of cohesive devices contributed to the success of the texts. This section analyses two TBRs according to the framework adapted from Hoey (1991). As was noted in Chapter 3 (§3.10.2), Hoey's description of different types of repetition actually covers a number of devices, grammatical and lexical, as defined by Halliday and Hasan (1976). For this reason, the term 'repetition' in this part of the study is applied in a wider sense according to the

typology posited by Hoey (1991) and summarized in §3.10.2. In his description, repetition includes pronominals that refer to the same person or thing. For example:

John lives in Yokohama. He has lived there for four years.

While repetition in Halliday and Hasan (1976) refers to the repetition of a particular word, repetition in Hoey (1991) refers to the reiteration of ideas. The *He* in the above sentence can be said to be a repetition or a reiteration of *John*, because it refers back to *John* in the previous sentence. Although the actual word is not repeated, the idea of *John* is carried into the next sentence. This is an important point to keep in mind in the analysis which follows, because there are indeed a number of different words that can refer to the same idea. For example, *John* may be referred to using *he*, *the man*, *my brother*, or *Mr Smith*.

We turn now to a closer analysis of the two TBRs, which are presented below:

Box 4.1: TBR 1 - A Beautiful Symbol of Friendship

1. Early in spring, cherry blossoms bloom along the Potomac River in Washington.
2. There are hundreds of trees along the river.
3. They look very beautiful, not only during the day but also at night.
4. Where did all these cherry trees come from?
5. Actually, they came from Japan!
6. They were a special gift from Japan to America.
7. At the beginning of the 20th century, the mayor of Tokyo sent the American people this wonderful gift from the Arakawa River.
8. He wanted to promote friendship between Japan and the United States.
9. Fifty years later, America returned the favor.
10. The old trees along the Arakawa River were dying, so America sent Tokyo many young cherry trees from the Potomac River.
11. This gift from America made the Arakawa River beautiful again.
12. Not only in Tokyo but also in Washington, everyone looks forward to the cherry blossoms every year.
13. These trees are still a beautiful symbol of Japanese-American friendship.

Box 4.2: TBR 2 - Abraham Lincoln

1. Abraham Lincoln was President of the U.S. during the Civil War between the northern and the southern states.
2. Lincoln was born on February 12, 1809, in a little log house and grew up on a farm.
3. The Lincoln family was very poor.
4. They were not able to send "Abe" to school or buy him books, but he borrowed books and read them again and again.
5. Abe studied hard and became a successful lawyer.
6. He was famous for his honesty.
7. They called him "Honest Abe."
8. One day, while he was still young, Abe saw something that he never forgot.
9. He was in New Orleans on business.
10. Some rich merchants were buying and selling black people from Africa.
11. Slavery was very common in the southern states in those days.
12. The slaves whom the white people bought worked on the farms in the South.
13. "I'll never see my children again!" one African woman was shouting.
14. "This man sold them to a rich white man who lives far from here."
15. The merchant who was selling the slaves saw Abe and asked, "You, sir, do you want to buy this woman?"
16. She works very hard."
17. "Of course not!" answered Abe.
18. "You can't sell people like dogs or potatoes!"
19. "Come on, come on!"
20. They're not people.
21. They're just slaves," said the merchant.
22. Lincoln was getting angrier and angrier.
23. "Someday, if I can, I'm going to stop this terrible business," Abe said.
24. The merchant laughed and continued his business.
25. Many years later, when Lincoln was President, he did not forget his promise.
26. In 1863, during the terrible Civil War, he kept his promise and freed all the slaves.

An analysis of these two TBRs revealed six important patterns and features, namely:

1. A variety of cohesive devices used in cohesive chains
2. The opening sentence and repetition links
3. The central topic or theme

4. Clusters of repetition
5. Sentences of varied length
6. The final sentence

These features are discussed below.

4.4.1 A variety of cohesive devices in cohesive chains

When the same idea is reiterated frequently in these TBRs, a variety of cohesive devices is used. For example, in TBR B, reiteration is achieved by the use of superordinates, reference, and direct repetition:

Table 4:11: Lexical chains and cohesive types in TBR B *A Beautiful Symbol Of friendship*

Sentence	Topic/idea	Cohesive type
1	Cherry blossoms	
2	trees	superordinate
3	they	reference, pronominal
4	tree	superordinate
5	they	reference, pronominal
6	they	reference, pronominal
10	trees	superordinate
12	cherry blossoms	synonym / hyponym
13	trees	superordinate

In Table 4.11 above, the same two devices are used consecutively only in sentences 5 and 6. In all other cases, especially in sentences 1 to 5, two different device types alternate. This adds variety to the text, while also offering new information. For example, a reader who knows nothing about cherry blossoms may learn that they are in fact trees.

In TBR 2, a similar variety of cohesive devices is used, namely repetition, superordinates, reference, synonym, and ellipsis.

Table 4:12: Lexical chains and cohesive types in TBR 2 *Abraham Lincoln*

Sentence	Idea/topic	Cohesive type
1	Abraham Lincoln	
2	Lincoln	repetition
3	Lincoln family	superordinate of <i>Abraham Lincoln</i>
4	They	reference – pronominal
5	Abe	synonym of <i>Abraham Lincoln</i>
6	He	reference – pronominal
7	him	reference – pronominal
8	he	reference – pronominal

The theme of *Abraham Lincoln* is reiterated in each of the first ten sentences. However, in each sentence a different word is used to achieve this repetition. Furthermore, despite constituting the smaller portion of cohesive devices in this study, superordinates, ellipsis and synonyms facilitate cohesion and improve the overall style of the text by breaking the monotony of using reference or repetition in every single sentence.

When two words are repeated verbatim, they are usually separated by a number of sentences. In TBR 1, *river* (sentences 2 and 11), *beautiful* (sentences 3 and 13), *Arakawa* (sentences 7 and 11) are separated by a number of sentences. Similarly, in TBR 2, *president* occurs in sentences 1 and 25, while *civil war* occurs in sentences 1 and 26. This suggests that direct repetition helps to create a direct link to the beginning of a written text. Moreover, since the links are separated by a number of sentences, redundancy and the monotony of over-repetition are avoided.

4.4.2 The opening sentence and repetition links

One of the primary functions of the opening sentence of a text is to introduce the main themes of the text. This is illustrated in the opening sentences of TBR 1 and TBR 2:

Early in spring, **cherry blossoms** bloom along the **Potomac River** in **Washington**. (TBR 1)

Abraham Lincoln was **President** of the **U.S.** during the **Civil War** between the northern and the **southern** states. (TBR 2)

The boxed words are the ideas or topics that are reiterated in the text. As can be seen, each sentence contains four or five such words (*Potomac, river, Washington* in TBR 1, and *President, Civil War, southern* in TBR 2). These are reiterated later either towards the middle (*river; southern*) or the end (*Potomac, Civil War*) of the text. This means that the opening sentences act as supporting frames from which the main ideas may be expanded. They introduce not only the main theme of the text (*cherry blossoms* in TBR 1 and *Abraham Lincoln* in TBR 2), but also ideas that are echoed at key points later on. They therefore do not exist in isolation, but play a vital role in ensuring that the most important ideas are introduced as early as possible in the text, and are then developed. Diagrams 4.3 and 4.4 below show the patterns of repetition in TBRs 1 and 2.

Diagram 4.1: Patterns of reiteration for TBR 1 *A Beautiful Symbol of Friendship*

Reiterated Type	Sentence Number												
	1	2	3	4	5	6	7	8	9	10	11	12	13
cherry blossoms	■	■	■	■	■	■				■		■	■
Potomac	■									■			
river	■	■					■				■		
Washington	■					■	■	■	■	■		■	■
beautiful			■										■
Japan					■	■	■	■		■		■	■
gift						■	■		■		■		
20th Century							■		■				
Arakawa							■				■		
mayor							■	■					
friendship								■					■

Diagram 4.2: Patterns of reiteration for TBR 2 *Abraham Lincoln*

	Reiterated Type	Sentence Number																										
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
1	Abraham Lincoln	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
2	President	■																										■
3	Civil War	■																										■
4	Southern	■																										
5	poor																											
6	family																											
7	books																											
8	lawyer																											
9	honesty																											
10	one day																											
11	saw																											
12	merchants																											
13	something (slavery)																											
14	bought																											
15	work																											
16	black																											
17	women																											
18	children																											
19	people																											
20	stop																											
21	promise																											

Note: The squares marked in black indicate an elided word.

Both Diagrams 4.3 and 4.4 have clear similarities. Firstly, in the top left-hand corner in both TBR 1 (Diagram 4.1) and TBR 2 (Diagram 4.2) an inverted L-shape can be seen. Sentence 1 indicates four ideas that are reiterated at different points of the text. The idea of *Abraham Lincoln* is reiterated throughout the text. The ideas of the *Southern* states are reiterated in the middle from sentences 9 to 12, while *President* and *Civil War* are reiterated in the final two sentences. The horizontal line that starts from the top left-hand corner of each diagram indicates a term or idea that is reiterated in the next few successive sentences, and then more sporadically toward the end of the text. Both texts contain a main topic (*cherry blossoms* in TBR 1 and *Abraham Lincoln* in TBR 2) that is reiterated throughout the text with some breaks. Secondly, both TBRs show a break from the main topic in successive sentences. In the case of TBR 1, the break occurs from sentences 7 to 9, while in TBR 2, the break can be seen from sentences 11 to 14. Thirdly, there are areas where a cluster of reiterated words occurs. In TBR 1, this can be seen around sentences 5 to 9 (*Japan, gift, 20th Century, Arakawa, mayor*), and in TBR 2, the biggest cluster occurs from sentences 10 to 16 (*bought, work, black, women, children, people*), with smaller clusters occurring

from sentences 2 to 6 (*poor, family, books, lawyer*) and 18 to 21 (*merchants, something, bought*). The similarities between these two texts do not appear to be an accident. Each text clearly shows (1) an opening sentence that contains links to the beginning, middle and end of the text; (2) a main topic (*cherry blossoms* in TBR 1 and *Abraham Lincoln* in TBR 2) formed by chains that are formed through reiteration, and (3) clusters of repetition where the shift in focus moves away from the main topic.

The precise meaning and significance of these three features are discussed in the next section.

4.4.3 The main topic or theme

The central topic or theme in the two texts can be identified by the way in which the ideas are reiterated in successive sentences throughout the text. A topic or idea that is reiterated most often, usually in successive sentences throughout the text, may be identified as the central topic. Ideas or topics that sustain chains for only a few sentences are seen as secondary topics.

In TBR 1 *cherry blossoms*, is reiterated in five consecutive sentences, while in TBR 2, *Abraham Lincoln* is reiterated in nine consecutive sentences. These are the main themes of the texts. The areas where the main topic does not occur (sentences 7–9 in TBR 1 and sentences 10–14 in TBR 2) indicate a shift in topic, and are discussed in §4.4.4 below. The main topics (*cherry blossoms* in TBR 1 and *Abraham Lincoln* in TBR 2) are then reiterated in some way in almost every sentence until the end of the text. This suggests that one way in which the text becomes coherent is through the sustained reiteration of the main topic, by means of different cohesive devices, throughout the text.

4.4.4 Clusters of repetition

Another noticeable similarity in the patterns of reiteration in Diagrams 4.3 and 4.4 is the presence of clusters of reiterated ideas. These clusters appear at various points in the texts. There may also be more than one such cluster. They may involve the main topic, but they may indicate a shift to other secondary topics. Since TBR 1 is a relatively short text, only one cluster occurs. It starts in sentence 5 (*Japan, gift, 20th century, Arakawa, mayor*) and continues until sentence 9. TBR 2 has one small cluster near the beginning from sentences 2 to 5 (*poor, family, books, lawyer*), a more prominent one from sentences 10 to 16 (*bought, work, black, women, children, people*) and a shorter one from sentences 18 to 21 (*merchants, something, bought*).

Clusters are an indication of elaboration on the main topic. In some cases, in order to elaborate fully, a complete shift in focus may occur such as in TBR 2, sentences 11 to 14, where the focus shifts from Abraham Lincoln to merchants selling slaves in the southern states. In other cases the elaboration is still firmly connected to the main topic, as happens in TBR 2 sentences 3 to 5 where the Lincoln family's poverty is described.

The cluster of cohesive devices in TBR 1 shows a complete move away from the main topic; in order to explain the origin of the cherry trees in Washington, the focus of the story moves away from the cherry trees to the mayor of Tokyo and his gift to America. The shift starts in sentence 5 (*Japan*) and continues to sentence 10, where the cherry blossoms are then reintroduced. In these sentences, the main focus becomes *Washington* and *Japan*, and the story of how the cherry trees came to Washington is told. The repetition clusters can be seen in Table 4.13 below.

Table 4:13: Repetition clusters in TBR 1 *A Beautiful Symbol of Friendship*

Sentence	Idea 1	Idea 2	Idea 3
5	They	Japan	
6		Japan	America
7		Tokyo	American
8		Japan	United States
9			America
10	trees	Tokyo	America
11	gift		
12	cherry blossoms		
13	trees		

In TBR 1, sentence 5 reads:

Actually, *they* (the cherry blossoms) come from *Japan*.

Considering the chain of reiteration in Table 4.13 above, *Japan* and *they* (cherry trees) co-occur in sentence 5. In sentence 6, the theme of Japan is continued, and the theme of America is introduced. Both themes are continued to sentence 8. In sentence 9, *America* is continued, while *Japan* falls away. In sentence 10, *America* continues the chain, while the secondary topics of *Tokyo (Japan)*, as well as the main topic (*trees*), are reintroduced and then reiterated in sentences 11, 12, and 13. The chain can therefore be said to be both vertical and horizontal: the horizontal chain consists of the sentences in which different secondary topics co-occur, while the vertical chain is the reiteration of ideas or topics.

Another example of a cluster which involves a shift in topic occurs in TBR 2. In sentence 9, the words *Abe saw something* is a cataphoric reference to the buying and selling described in sentence 10. A new secondary topic is introduced at this juncture, and developed with the introduction of the topics of *merchant, slavery and black buying and selling*. There also

appears to be no link between sentences 9 and 10. However, if one bears in mind that sentence 10 is describing what Lincoln saw, it becomes clear that the phrase *Abe saw that* has been elided, and the elided *he saw* then creates a link to sentence 10.

What follows from sentences 10 to 14 appears to be a complete shift. However, it is all related to Lincoln, because it describes what he saw. From sentence 10 to 14, there is no mention of Abraham Lincoln, while the main focus becomes *slavery* (selling) (sentences 10 to 12), and *buying and selling* (sentences 12–15). Again, this cluster provides elaboration on the story, telling the reader about the slave trade in the American South. The main ideas are reiterated through the words *merchant*, *slavery*, *bought*, *work*, *black*, *woman*, and *children*. The actual cohesive chains used to create this cluster are now examined.

Table 4:14: Repetition clusters (2) in TBR 2 *Abraham Lincoln*

Sentence	Idea 1	Idea 2	Idea 3	Idea 4
8	something (cataphoric reference)			
9	He			
10	He	buying and selling		
11		slavery		
12		slaves		
13		African	woman	
14			(she said)	this man
15	Abe	slaves		merchant

In Table 4.14, the boxed areas show how the topics shift from Abe to slavery in sentence 10, and then to the *slave woman* in sentence 13. The slave woman then refers to one of the merchants in sentence 14. The *merchant* topic is then carried into sentence 15, where the main topic, *Abraham Lincoln*, is reintroduced. This shows that a shift in topic usually starts

with a sentence that contains the original topic and the new topic. For example, sentence 10 contains both *he* (implied), and *buying and selling*.

The two cases above describe a complete shift in topic. However, there are also cases in which clusters indicate elaboration without such a shift. This can be seen in TBR 2, where the small cluster in sentences 3, 4 and 5 does not shift away from the main topic, but rather elaborates on Lincoln's situation as a child, using the ideas of *poor, family, read, lawyer*

Table 4:15: Repetition clusters (1) in TBR 2 *Abraham Lincoln*

Sentence	Idea 1	Idea 2	Idea 3	Idea 4	Idea 5
3	Lincoln	family	poor		
4	Abe	They	unable to buy	books	
5	Abe			studied	lawyer

In the cluster in Table 4.15 above, the new information is still connected to the main topic. Another important point is that only one idea (*Abe*) is repeated directly. All other ideas are repeated using hyponyms (*Lincoln – Abe*). It can be safely assumed that a person's given name is essentially a hyponym for their family name. Other reiterations are *family – they* (pronominal), *poor – unable*, *books – studied* (collocation). The word *lawyer* is not repeated in this cluster, but is later reiterated with the collocate *business* in sentence 9. A notable feature of this cluster is the way in which sentence 4 builds on sentence 3 with the use of *Abe, they (the Lincoln family)*, and *unable to buy (poor)*, and with the addition of *books*. Sentence 5 is still rooted in the main topic (*Abe*), but then also reiterates the idea of *books (studied)*, and then elaborates further with the word *lawyer*.

This cluster shows how information can be built up incrementally, but is always anchored in the main topic. Even when a shift does occur, such as from Abraham Lincoln to the merchants in TBR 2, these clusters do not appear completely in isolation from the main

topic. They are intrinsically related to the main theme, but also expand on it. In both TBRs the clusters are, at both their beginning and their ending, connected to the main topic.

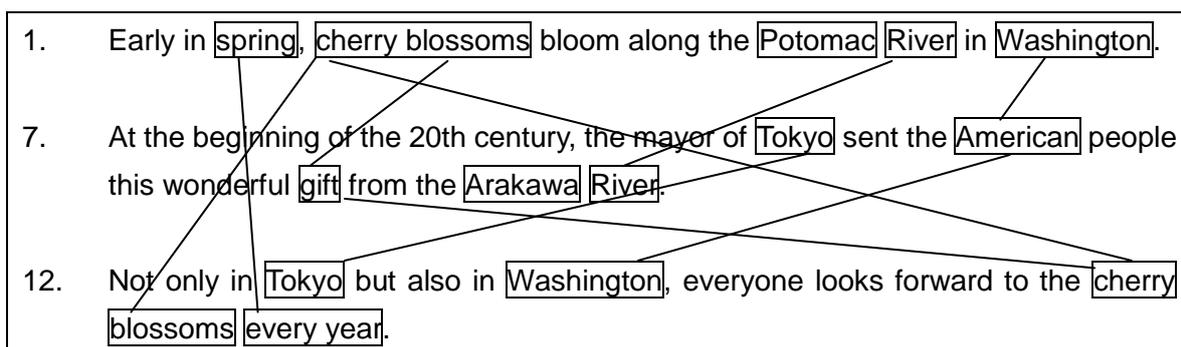
4.4.5 Sentences of varying length

A text that contains sentences of varying lengths has two distinct advantages, namely, the creations of bonds in long sentences and the contrast of long with short sentences.

As already noted (§3.10.2), when the same idea occurs in two sentences, a link is formed between these two sentences. If three or more ideas are reiterated in the same two sentences, a bond is formed. Since longer sentences by their very nature contain more lexical items, the chances of forming bonds with other sentences naturally increase.

As Hoey (1991) has noted, it is often difficult, if not impossible, to create repetition bonds in a narrative text to the extent found in expository texts, largely because in narrative texts much of the cohesion is achieved in other ways such as references to time. Nevertheless, these references are clearly illustrated in TBR 1, where bonds exist between sentences 1 and 7 and between sentences 7 and 12. These bonds are illustrated in Diagram 4.3 below.

Diagram 4.3: Bonding sentences in TBR 1 *A Beautiful Symbol of Friendship*



In Diagram 4.3, sentences 1 and 7 contain three links, as do sentences 7 and 12.

Sentences 1 and 7 are linked by

cherry blossoms	gift (superordinate)
River	river (direct repetition)
Washington	American (superordinate)

Sentences 7 and 12 are linked by the following:

<i>Tokyo</i>	<i>Tokyo</i> (direct repetition)
<i>American</i>	<i>Washington</i> (hyponym)
<i>gift</i>	<i>cherry blossoms</i> (hyponym)

Not only do the three links between these pairs of sentences create bonds, but when read together these three sentences combine to form a coherent text, despite being separated by a number of sentences in the text. These three sentences thus combine to form the central idea or theme of the text.

The creation of bonds between sentences, while not necessarily vital in narrative texts (TBR 2 does not in fact contain such bonded sentences), nevertheless means that the key ideas and themes are semantically linked to one another, thus improving textual coherence.

However, it is not only long sentences that are important. Shorter sentences are also very useful in adding more impact to the text. This can be seen in TBR 2, sentences 2 to 4, which are provided below:

2. Lincoln was born on February 12, 1809 in a little log house and grew up on a farm.
3. The Lincoln family was very poor.
4. They were not able to send "Abe" to school or buy him books, but he borrowed books and read them again and again.

Sentence 2 is relatively long, giving information on when and where Lincoln was born, and where he grew up. However, sentence 3 is very short. By using a short sentence after a long sentence the writer increases the impact of the information by creating a contrast. In this case, the shorter sentence also marks a secondary topic sentence which is then immediately elaborated on in sentence 4 by describing the family's inability to buy books.

What these two examples illustrate is that it is the combination of the elaboration and links to other parts of the text in longer sentences, together with the impact of the shorter sentences, that enhance the success of a text.

4.4.6 A final sentence that links back to the middle and beginning

In both TBR 1 and TBR 2, the final sentences contain a number of ideas that are linked to the beginning, the middle and the end of the texts.

In TBR 1, in the final sentence

13. These trees are still a beautiful symbol of Japanese-American friendship.

five words (*trees, American, beautiful, Japanese, friendship*) are all reiterations of ideas that occur earlier in the text. This sentence also contains one final mention of the main topic (*trees*) that runs through the text, as described in §4.4.1. Other words in the sentence also refer back to the middle of the text (*friendship* appears in sentence 8), as well as back to the beginning (*beautiful* in sentence 3). This can be seen clearly in Diagram 4.1 in §4.4.2.

Similarly, as can be seen in Diagram 4.2 in §4.4.2, the final sentence in TBR 2, contains four ideas (*Lincoln, Civil War, promise, slaves*) that are linked to previous sentences in different parts of the text: *Abraham Lincoln* is the main topic – thus it runs through the text even though it is not referred to in each sentence. *Civil War* links directly back to sentence

1. The idea of a *promise* is linked to sentence 24. The idea of *slaves* is developed in the text from sentence 8 to sentence 21.

Of course, it is possible to have a short final sentence. In that case, the penultimate sentence may perform the function of summarizing and drawing together the main themes of the text. Such a sentence performs two important functions. Firstly, it reminds the reader of the salient facts, namely, Abraham Lincoln, his encounter with the slave trade, and how he decided to stop it, and secondly helps to give the text a sense of wholeness and completeness.

Having examined these TBRs and having found six features of cohesion in their patterns of repetition, the next step was to determine the extent to which the STs also contained these features.

4.5 FINDINGS RELATING TO QUESTION 4

This section addresses the finding relating to Question 4: How do the patterns of reiteration in the STs compare to those in the TBRs?

In order to address this question, six STs, two graded high, two graded average, and two graded below average were analyzed.

4.5.1 Student texts: a note on grading

The results of the test of inter-rater reliability described in Chapter 3 (§3.9) are presented in Table 4.16 below:

Table 4:16: Rater scores for the six selected student texts

	Researcher	Rater 1	Rater 2	Rater 3	Rater 4
Tom Sawyer	5	5	5	5	5
Jumble	4	5	5	4	5
Phantom	3	4	3	3	3
Escape	3	3	3	3	3
Scarf	2	3	2	2	2
Sangokushi	1	2	1	1	3

The inter-rater correlation was 0.91 (see §3.9), which suggests a high agreement among the five raters (me and four colleagues). Despite the small sample size and the limited scale of one to five, the correlation does reflect a high degree of consensus on a level of overall impression as to what constitutes a strong or a weak text.

4.5.2 Criteria for analysis

As was noted in §4.4 above, analysis of the TBRs revealed six important criteria for successful cohesion in a text. I then examined each of the STs to determine how they compared to the TBRs in the light of these criteria:

1. A variety of cohesive devices used in cohesive chains
2. The opening sentence and repetition links
3. The central topic or theme
4. Clusters of repetition
5. Sentences of varied length
6. The final sentence

4.5.3 Results of the analysis

For the sake of simplicity, analyzed STs were numbered ST 1 to 6. This numbering system differs from the one used in the quantitative analysis.

Very good to excellent:

- ST1 Tom Sawyer
- ST2 Jumble Sale

Average:

- ST3 Escape
- ST4 Phantom of the Opera

Below average:

- ST5 The Scarf
- ST6 Sangokushi

The following sections analyze each ST according to the criteria set out in §4.5.2.

4.5.4 ST 1 Tom Sawyer

This section examines ST1 (*Tom Sawyer*) in the light of the features identified in the TBR analysis. It looks first at what cohesive devices are used.

Box 4.3: Text for ST 1 *Tom Sawyer* 93 words: grade – very good

1. "The adventures of Tom Sawyer" is a story about a boy whose name is Tom Sawyer.
2. Tom Sawyer who liked playing and having adventures lived with his Aunt Polly, because his parents were dead.
3. On Friday he didn't go to school and went to the river.
4. Polly was angry and said, "Tomorrow you have to paint the fence."
5. Saturday morning Tom was getting sadder and sadder, but he started to paint.
6. This work was hard but he wanted to finish quickly.
7. Then what did he do?
8. There are many chapters in this book and I like this story best.
9. Tom painted the fence happily and he made his friends believe the painting was interesting.
10. Many wanted to paint and some of them even gave him presents.
11. I think Tom is a very clever boy.

4.5.4.1 A variety of cohesive devices used in cohesive chains

Table 4.17 shows how the writer uses a number of cohesive devices in order to reiterate the idea of *Tom Sawyer*, which is the main topic of the sentence.

Table 4:17: Lexical chains and cohesive types (1) in ST 3 *Tom Sawyer*

Sentence	Topic	Cohesive devices
1	Tom Sawyer	
2	Tom Sawyer	direct repetition
3	You	reference – pronominal
4	Tom	hyponym
5	Tom	repetition
6	He	reference – pronominal
7	He	direct repetition
8		
9	Tom	
10	him	reference – pronominal
11	Tom	direct repetition – Tom

Since this topic is reiterated in every sentence except sentence 8, it is clearly the main topic of the text. As can be seen, three different cohesive devices (direct repetition, reference and hyponyms) are used. Although direct repetition is used three times (sentences 2, 4 and 7), on each occasion a different word (*Tom Sawyer*; *Tom*; *He*) is repeated. Moreover none of these repeated words occur more than twice consecutively. Instead, different devices are used interchangeably, and direct repetition is interspersed with other devices such as hyponyms and reference. For example, in the first six sentences *Tom Sawyer* is reiterated as follows: *Tom Sawyer – Tom Sawyer – You – Tom – Tom – He*. Here one can see the variety of devices used to refer to Tom Sawyer.

Another cohesive chain can be seen from sentences 4 to 6, using the secondary topic *paint the fence* (sentence 4), which in itself refers to and elaborates on the main topic (*Tom Sawyer*).

Table 4:18: Lexical chains and cohesive types (2) in ST 3 *Tom Sawyer*

Sentence	Idea/topic	Cohesive type
4	paint the fence	
5	paint	superordinate
6	work	superordinate

In Table 4.18, although *paint the fence* in sentence 4 consists of three tokens, it essentially constitutes one idea. In this sense, the word *paint* in sentence 5 may be considered a superordinate of *paint the fence*, since painting may refer either to painting a picture, or covering something (in this case a fence) in paint. In sentence 6, the writer uses yet another superordinate (*work*) to refer to painting the fence. This demonstrates a thorough understanding of how to use superordinates in order to enhance a text. Not only do superordinates offer a wider explanation of the words to which they refer, but they also add variety by repeating the idea but not the word itself. This helps to break the monotony so often found in weaker texts.

Ellipsis also occurs in sentence 10:

9. Tom painted the fence happily and he made his friends believe the painting was interesting.
10. Many (friends) wanted to paint and some of them even gave him presents.

In sentence 10, *friends* is elided, which suggests a slightly more sophisticated understanding of how some direct repetition may be unnecessary. This suggests that the writer knows when words should not be repeated unnecessarily.

Finally, the use of words of emotion (*angry* – sentence 5; *sadder* – sentence 6; *happily* – sentence 9) also sets up reiteration links using antonyms. This helps to reinforce the links between the sentences, while also providing more important information. It also suggests

that the writer is not simply telling the story and offering his point of view, but also trying to offer some kind of interpretation of the story that he read.

4.5.4.2 The opening sentence and repetition links

Diagram 4.4 shows that the opening sentence contains four ideas that are reiterated. The opening sentence reads as follows:

Diagram 4.4: Patterns of reiteration for ST 1 *Tom Sawyer*

	Reiterated Type	Sentence Number										
		1	2	3	4	5	6	7	8	9	10	11
1	Tom Sawyer	■	■	■	■	■	■	■		■	■	■
2	adventures	■	■									
3	Story	■							■			
4	boy	■										■
5	Aunt Polly		■		■							
6	Friday			■	■	■						
7	paint fence				■	■	■			■	■	
8	angry				■	■				■		
9	hard						■			■		
10	friends									■	■	

The first sentence reads:

"The adventures of Tom Sawyer" is a story about a boy whose name is Tom Sawyer.

The boxed words are those which are reiterated in the text. Apart from *Tom Sawyer*, which was dealt with in the previous section, three other ideas in this opening sentence are repeated in the text:

- adventures* reiterated in sentence 2 (direct repetition – *adventure*)
- story* reiterated in sentence 8 (synonym – *book*)
- boy* reiterated in sentence 11 (direct repetition – *boy*)

This means that, while the main idea, *Tom Sawyer*, runs through the text, these three additional ideas create links between the opening sentence and the beginning (in the idea of *adventures*), the middle (*story or book*) and the end (*boy*) of the text. Thus, the pattern found in this ST is somewhat similar to those found in the TBRs. First, there is the inverted L-shape, secondly, there is the main topic which runs through the text, and thirdly, there is a cluster of repetition. This suggests that successful student texts exhibit similar patterns.

4.5.4.3 The central topic or theme

As can be seen in Diagram 4.4, *Tom Sawyer* is the main topic of this text. It was noted above that reiteration of *Tom Sawyer* occurred in 10 of the 11 sentences in the text.

4.5.4.4 Clusters of repetition

Two clusters can be found in this text. The biggest cluster can be seen in Diagram 4.4 above and occurs in sentences 2 to 6. A smaller cluster occurs in sentences 9 and 10.

The bigger cluster (sentences 2 to 6) is anchored by the main topic (which is *Tom*), but helps to elaborate more on the situation regarding Tom being ordered by Aunt Polly to paint the fence. The repetition chains can be seen in Table 4.19 below.

Table 4:19: Repetition clusters in ST 1 *Tom Sawyer*

Sentence	Idea 1	Idea 2	Idea 3	Idea 4	Idea 5
2	Tom Sawyer	Aunt Polly			
3	Tom Sawyer		Friday		
4	You	Polly	Tomorrow	paint the fence	
5	Tom		Saturday	paint	sadder
6	He			work	

While the main topic (*Tom Sawyer*) is sustained through this phase of the text, two other cohesive chains, namely, time markers (*Friday, tomorrow, Saturday*) and the theme of paint, continue for three sentences through the text.

The second smaller cluster in sentences 9 and 10 explains how Tom Sawyer makes his friends believe that what they are doing is actually fun. As noted in §4.5.4.1 above, there are a number of links between these two sentences:

- 9 Tom painted friends
- 10 him paint Many (friends elided)

This cluster is still rooted in the main topic, but further elaborates on and explains the story.

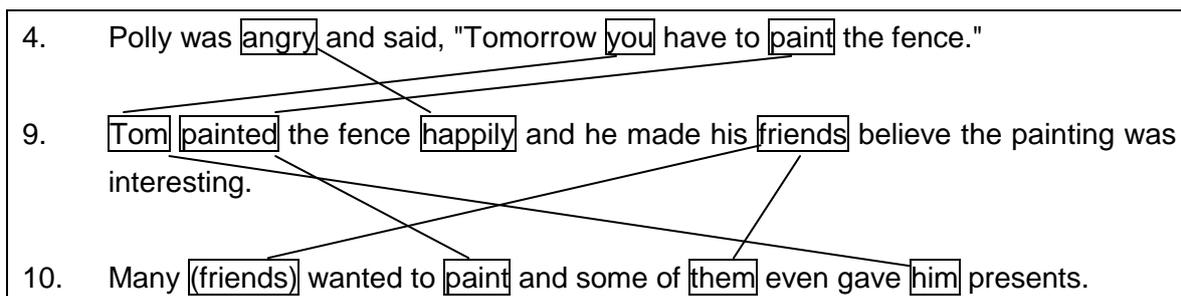
4.5.4.5 Sentences of varied length

In this essay, there is a variety of long and short sentences.

Sentences 1 and 2 both use subordinate clauses (*whose name is Tom Sawyer* in sentence 1 and *because his mother is dead* in sentence 2). Sentences 3, 4, 5, 8 and 9 are all compound sentences using either *and* or *but* as conjunctions. Sentences 7 and 11 are both simple sentences that provide pauses between longer sentences, and add more impact to the text.

The presence of longer sentences also allows for the formation of bonds between sentences 4 and 9, and sentences 9 and 10:

Diagram 4.5: Bonded sentences 4, 9 and 10 in ST 1 *Tom Sawyer*



As can be seen in Diagram 4.5, the bonded sentences create a good summary of the story, even though sentences 4 and 9 are separated by four sentences.

4.5.4.6 The final sentence

In this ST, the final sentence is rather short, so it contains very few links. However, combined with the penultimate sentence more links are established.

The final sentence itself only links back to the beginning of the text:

1. I think Tom is a very clever boy.

The links are as follows:

Tom – Tom Sawyer sentence 1
 Boy – boy sentence 1

Although the final sentence does not fit the pattern of the TBRs, the penultimate sentence contains three links to different parts of the text:

10. Many wanted to paint and some of them even gave him presents.

The links can be seen in Table 4.20 below:

Table 4:20: Lexical links to the penultimate sentence in ST 1 *Tom Sawyer*

Sentence	Idea 1	Idea 2	Idea 3	Idea 4	Idea 5
1	Tom Sawyer (book title)		story		
2	Tom Sawyer				
3	Tom Sawyer				
4	You (Tom)		paint the fence		
5	Tom		paint		
6	He		work		
7	he				
8		story			
9	Tom		painted the fence		friends
10	him		paint	them	

Despite being the penultimate sentence, sentence 10 performs the same role as the final sentences in the TBRs by creating links to the beginning (*Tom*), middle (*paint the fence*) and end (*friends*) of the text. This sentence, combined with the final sentence, functions to summarize the text's main themes and ideas.

4.5.4.7 Summary

The extent to which ST 1 matches the features identified in the TBRs can be summarized as follows:

1. It makes use of a variety of cohesive devices.
2. It contains an opening sentence with links to different parts of the text.
3. A central topic or theme is developed throughout the text (*Tom Sawyer*).
4. Two clusters of repetition, one that elaborates on the main topic and another that demonstrates a shift to another topic, were identified.
5. Sentences vary in length.
6. Although the final sentence is short, the penultimate sentence contains links to the beginning, the middle and the end of the text.

4.5.5 ST 2 The Jumble Sale

In this section ST 2 (*The Jumble Sale*) is discussed.

Box 4.4: Text for ST 2 *The Jumble Sale* 110 words: grade – very good

1. 'The jumble sale' is a story about old toys that the children liked.
2. Mom and Dad and their children were cleaning their house.
3. Dad threw useless things out.
4. The old toys that the children liked threw out.
5. And these useless things went to the jumble sale.
6. The children were getting sadder and sadder.
7. Then what happened after it?
8. This story is the most interesting book in English that I read for me because the children were as sad as me.



9. I was sad when the toys that I liked were broken, too.
10. So I can imagine and understand the children's mind.
11. So I like it and would recommend this book.
12. I can imagine and understand the children's mind, because I was (also) sad when the toys that I liked were broken,

4.5.5.1 A variety of cohesive devices used in cohesive chains

A number of cohesive chains are found in this text. The most important of these are presented in Table 4.21 below:

Table 4:21: Lexical chains and cohesive types (1) in ST 2 *The jumble sale*

Sentence	Idea/topic	Cohesive type
1	'Jumble'	
2	useless things	general item
3	old toys	synonym (hyponym for <i>useless things</i>)
4	jumble	synonym - <i>old toys</i>
9	broken toys	synonym - <i>jumble</i>

Each time this writer refers to the toys that are to be thrown out, he uses a different cohesive device, for example *jumble* – *useless things* – *old toys* – *broken toys*. This is an excellent example of reiteration that adds interest and new information to the text. This suggests that the writer has a solid grasp of the importance of adding variety to the text. He is also able to use the model text that I provided, as well as the book being reviewed.

A second chain, illustrated in Table 4.22, occurs from sentences 6 to 9.

Table 4:22: Lexical chains and cohesive types (2) in ST 2 *The Jumble Sale*

Sentence	Idea/topic	Cohesive type
6	sadder	
7	it	pronominal (should be demonstrative <i>this</i>)
8	sad	repetition – <i>sadder</i>
9	sad	repetition – <i>sad</i>
10	mind	superordinate – <i>sad</i>

While learners at this level may be expected to use cohesive devices to refer to particular words or items, this writer displays a far deeper understanding of how cohesion can be used. The use of *it* in sentence 7, rather than simply referring to a particular word or idea, refers to an entire incident of the toys being sold and the children becoming sadder. This reveals a much higher level of awareness than most of the learners in the group. Moreover, although it should actually be *this*, the use of *it* in this case helps to break the monotony of the word *sad* being repeated four times. Finally, the use of the superordinate *mind* reiterates the children's sadness. This is achieved by referring to a larger category of state of mind, rather than by simply repeating the children's emotion.

The third chain (Table 4.23), while broken in places, contains another important theme in the text, namely, that of *the children*:

Table 4:23: Lexical chains and cohesive types (3) in ST 3 *The Jumble Sale*

Sentence	Idea/topic	Cohesive type
1	children	
2	children	direct repetition
4	children	direct repetition
6	children	direct repetition
8	children	direct repetition
10	children	direct repetition

Again, the monotony of repeating the word *children* is broken by the word not being repeated always in successive sentences. In fact, the only consecutive sentences in which the word is repeated are sentences 1 and 2. Otherwise, the theme is repeated in every second sentence. While this is not a direct chain, it still creates links with six out of ten sentences in the story.

4.5.5.2 The opening sentence and repetition links

The Jumble Sale is a story about old toys that the children liked.

The opening sentence has four ideas that are reiterated in the text:

<i>Jumble</i>	reiterated in sentences 2, 3, 4 and 9 (also using old toys)
<i>story</i>	reiterated in sentences 8 and 11 (<i>story, it</i>)
<i>children</i>	reiterated in sentences 2, 4, 6, 8, 10 (the children)
<i>liked</i>	reiterated in sentences 4 and 9 (<i>liked</i>)

As in the TBRs and ST 1 *Tom Sawyer* above, the inverted L can be clearly seen in Diagram 4.6, albeit with one token missing. The fact that *jumble* does not occur in the second sentence is made up for by the repetition of *children* in sentences 1 and 2. Thus, some known information is carried from sentence 1 to sentence 2. The theme of *jumble* or *old toys* is then carried through to sentence 5 by the words *useless things* and *toys*.

Diagram 4.6: Patterns of reiteration for *The Jumble Sale*

Jumble Sale												
Reiterated Type		Sentence Number										
		1	2	3	4	5	6	7	8	9	10	11
1	jumble	■		■	■	■			■	■		
2	sale	■				■						
3	story	■							■			■
4	toys	■			■					■		
5	children	■	■		■		■		■		■	
6	liked	■			■					■		
7	Mom and Dad		■	■								
8	cleaning / threw out		■	■	■							
9	get sadder						■	■	■	■	■	
10	interesting								■			■

4.5.5.3 The central topic or theme

At first glance, the main theme of the text appears to be that of the *jumble*. However, it is not continued through the entire text. In fact, its last reiteration is in sentence 5. On the other hand, the theme of the *children* is reiterated throughout the text, albeit not in every sentence. Thus, *children* could be considered to be the main theme of the text.

4.5.5.4 Clusters of repetition

Two clusters can be found in this text. The first appears in sentences 2 to 5, and the second in sentences 8 to 10. The nature and meaning of these clusters is discussed below.

The cluster in sentences 2 to 5 contains chains that are illustrated in Table 4.24.

Table 4:24: Repetition clusters (1) in ST 2 *The Jumble Sale*

Sentence	Idea 1	Idea 2	Idea 3	Idea 4
2	children	Mom and Dad	cleaning	
3		Mom and Dad	threw out	useless things
4	children		threw out	old toys
5				useless things

In this cluster, *throwing out* collocates with *cleaning*. This theme of cleaning indicates a brief shift in topic. As in the previous texts analyzed, words are linked either by their presence in the same sentence, or by their reiteration in subsequent sentences. As a result, sentences 2 and 4 contain two of the same ideas that form the links, namely, *children* and *cleaning/throw out*.

It may appear as if the above sentences are being over-repetitive. However, each sentence also contains important new information. In sentence 2, the new information is *cleaning the house*. In sentence 3 it is *throwing useless things out*. In sentence 4, we are told that these things were actually liked by the children. Finally, in sentence 5, we are told that discarded items were not actually thrown out, but went to the jumble sale mentioned in the first sentence. Thus, not only is new information introduced, but it is always anchored in known information. This means that the writer is not simply repeating information, but developing the main topic through elaboration. By providing a variety of cohesive devices to describe the jumble, he is not only reiterating its uselessness, but also offering some comment on the toys. It is this development and elaboration on a topic or theme that enhances the success of a text – even if some grammatical mistakes may be present – and helps to maintain reader interest in the story. It is largely for these reasons that the text obtained a high grade.

The second cluster occurs from sentences 8 to 10.

Table 4:25: Repetition clusters (2) in ST2 *The Jumble Sale*

Sentence	Idea 1	Idea 2	Idea 3	Idea 4	Idea 5	Idea 6
8	story	children	sadder			
9			sad	jumble	toys	liked
10		children	mind			

As can be seen in Table 4.25, the main theme in this cluster is *sad*. This creates links with all three sentences. The reason for the link with *mind* has to do with some L1 interference. In Japanese, the word *kokoro* is used to refer both to one's emotions and one's thoughts, or mind. Thus, while to a native English speaker, the word *mind* may seem to refer to thinking, in this case it refers to the children's emotions, which were sad. The link between sentences eight and nine allows for the important elaboration in sentence 9 where the writer explains his own emotional reaction to the book. It is also worth noting that, while offering his own reaction and interpretation of the text – one of the criteria for a higher grade – he reiterates the important theme of the toys that the children liked:

8. This story is the most interesting book in English that I read for me because the children were as sad as me.
9. I was sad when the toys that I liked were broken, too.
10. So I can imagine and understand the children's mind.

As a result, although he is offering new information, he is simultaneously reinforcing the main themes of the text.

4.5.5.5 Sentences of varied length

This text contains 11 sentences, six of which are simple sentences (sentences 2, 3, 5, 7, 10), four complex sentences with relative clauses (sentences 1, 4, 8, 9), and one a compound sentence (sentence 11). Since the learners had only recently been introduced to relative clauses (about three weeks prior to the writing of these essays), this learner shows a grasp of how the relative clauses could be used to present more information in a sentence. One could argue that this learner should have combined one or two simple sentences (for example, sentences 3 and 4, or sentences 5 and 6) so as to make the text flow better. However, this should be seen more as the next step forward in his development as a writer, rather than any shortcomings he possessed at this particular level.

The presence of complex and compound sentences also allows for a bond to form according to Hoey's (1991) criteria:

Diagram 4.7: Bonded sentences 1, 4 and 9 in ST 2 *The Jumble Sale*

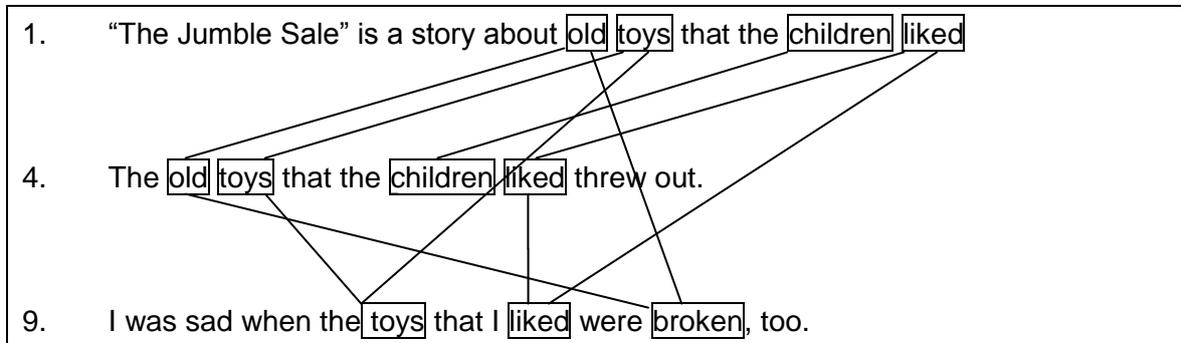
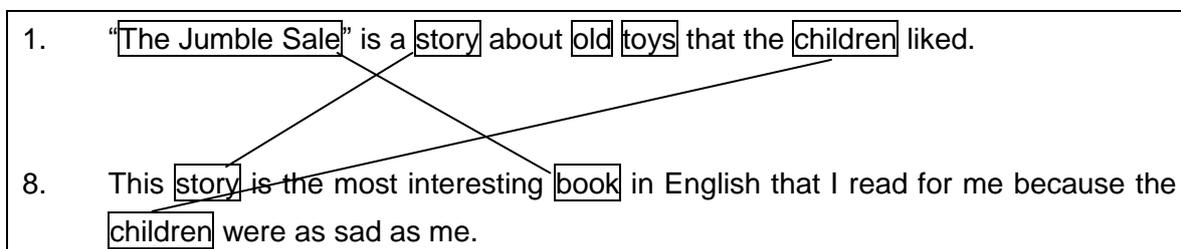


Diagram 4.7 demonstrates the links that exist between sentences 1, 4 and 9, creating bonds between the three sentences. The repetitiveness of the three sentences as presented here is not apparent in the main text, since two or three sentences separate each of them. Moreover, when read together the three sentences form a cohesive whole. They constitute a summary of a book review, with sentence 1 acting as an introduction, sentence 4 presenting the problem phase of a story, and sentence 9 presenting the writer's impression of the story.

The bond between sentences 1 and 8 is illustrated in Diagram 4.8.

Diagram 4.8: Bonded sentences 1 and 8 in *The Jumble Sale*



These two sentences read together form not simply a bond, but also constitute a coherent text, essentially a summary of the entire text.

As in the previous texts, the key to the formation of these bonds lies in the use of adjectival phrases and subordinate clauses which, in providing more information, allow for the reiteration of this information in subsequent sentences. In sentence 1, four of the six words that formed links with other sentences were found either in the adjectival phrase (*about the old toys*), or in the relative clause (*that the children liked*).

In addition to adjectival phrases, complex sentences and compound sentences, this writer is able to use comparatives (*the children were as sad as me* – sentence 8), another grammatical structure that was learned only a few weeks prior to the test in which this essay was written. This suggests that the writer of the text was quickly able not only to grasp the concept of when and how compound sentences are constructed, but also to apply this knowledge in an extended text.

4.5.5.6 The final sentence

The final two sentences read as follows:

10. So I can imagine and understand the children's mind.
11. So I like it and would recommend this book.

While linking back to the middle and beginning of the text by means of the word *story* in sentences 1 and 8, the final sentences constitute one of the very few weaker points of this otherwise successful text. The final sentence contains the following links, which are illustrated in Table 4.26:

Table 4.26: Links to the penultimate and ultimate sentences in ST 2 *The Jumble Sale*

Sentence	Topic/idea	Cohesive type
1	story	
8	story	(repetition)
11	book	(synonym)
Sentence	Topic/idea	Cohesive type
8	interesting	
11	like	(collocation) with <i>story</i> (sentence 8)

However, as in ST 1 (*Tom Sawyer*), it is in fact the penultimate sentence that contains important information that links back to the beginning and middle of the text:

10. So I can imagine and understand the *children's mind*.

As shown in §4.5.5.1, *children* appeared in sentences 1, 2, 4, 6, 8 and 10, while *mind* is a superordinate of *sad*, which previously appeared in sentences 6, 7, 8 and 9.

This suggests that one sentence formed from these last two sentences would have created a greater number of links to different parts of the text, thus offering a more complete summary of the text. For example:

I can imagine and understand the children's minds, because I was (also) sad when the toys that I liked were broken.

A sentence such as the one above would have contained links to the main topic (*children*), as well as to the secondary topics (*sad*, *toys*, *broken*), thus also connecting the final sentence with the middle and beginning of the text. In this way, a stronger conclusion to this text could have been achieved.

4.5.5.7 Summary

ST 2 *The Jumble Sale* contains all six features found in the TBRs. However, as in ST 1, the penultimate sentence rather than the final sentence is the one that serves to summarize the text by creating links to the beginning, middle and end. This sentence uses a wide variety of cohesive devices to link the sentences together. Furthermore, it elaborates on the main themes (namely, the *children* and the *jumble*) by expanding on the known information. This helps to maintain reader interest in the story and is achieved by the use of complex and

compound sentences, as well as shorter sentences that help to break the monotony of the text. For all these reasons, the text was able to achieve a high grade.

4.5.6 ST 3 Escape

Box 4.5 ST 3 *Escape* 64 words: grade – average

1. An innocent man that is in the prison decides to escape from there like a spider.
2. While he runs away, he steals a bicycle and someone's clothes.
3. He becomes a true thief.
4. He is caught and returns to prison.
5. Then he have got an idea when he sees a mole in this time.
6. The story finishes here.
7. What will happen if the story continues?

One of the main weaknesses with this particular text is structural: it does not introduce the book as a review, nor does the writer offer any personal impression of the book. Had the writer done so, he may well have increased the number of links between sentences and created a more successful text in the process. However, since the goal of this analysis is to determine what the writer actually did do, the focus now turns to some of the chains and the variety of devices that the writer used.

4.5.6.1 A variety of cohesive devices used in cohesive chains

The longest chain of reiteration in the text is the main topic of the *man* (sentence 1). This theme (indicated in Table 4.27) is continued for the next four sentences.

Table 4:27: Lexical chains and cohesive types (1) in ST 3 *Escape*

Sentence	Topic/idea	Cohesive type
1	man	
2	he	reference pronominal
3	he	repetition
4	he	repetition
5	he	repetition

As can be seen, this particular chain contains four repetitions of the word *he*. The writer does not seem to be aware of other cohesive possibilities, such as the character's name, or perhaps a superordinate such as *innocent man*, or *this thief*.

The second chain is a little more sophisticated:

Table 4:28: Lexical chains and cohesive types (2) in ST 3 *Escape*

Sentence	Topic/idea	Cohesive type
1	innocent	
2	steals	collocation – <i>prison</i> – sentence 1
3	thief	collocation – <i>steals</i>

In Table 4.28, it could be argued that *thief* is an antonym of *innocent*. In this chain, the writer makes good use of collocation, for example the use of *prison*, *innocent*, *escape*, *steals*, *thief*. However, the ideas in this particular chain are not reiterated or reinforced later in the text.

The third chain uses a different device in each reiteration (namely, a synonym and an antonym), but is too short to have any lasting impact on the text:

Table 4:29: Lexical chains and cohesive types (3) in ST 3 *Escape*

Sentence	Topic/idea	Cohesive type
1	escape	
2	runs away	synonym
3		
4	caught	antonym

Again, as can be seen in Table 4.29, the learner uses the passive *caught*, which is a structure that has not been taught yet in our class. How he knows how to use this is open to speculation: he may have used the textbook, or he may actually know this phrase through

some previous study, or he may have received help from outside. Be that as it may, this chain is very effective in relating the main points of the story. However, it is not linked to later sentences.

Finally, in the last two sentences, the writer makes use of antonyms (*stop – continue*) to awaken the reader's interest in the story. Here, the writer demonstrates his knowledge of the vocabulary (*continue*) that has been recently introduced in the textbook.

4.5.6.2 The opening sentence and repetition links

An innocent man that is in the prison decides to escape from there like a spider.

The opening sentence is, despite a lack of introduction to the book itself, a very strong sentence, containing a number of words that are reiterated in the text. In addition to *innocent*, *man* and *escape*, all of which were dealt with in the previous section, *prison* is reiterated in sentence 4. The word *spider* can be said to be reiterated by *mole* in sentence 5, through the theme of getting his ideas from the small animals that he sees. This is reinforced by the direct repetition of *like a* in sentence 5.

Diagram 4.9: Patterns of reiteration for *Escape*

	Reiterated Type	Sentence Number						
		1	2	3	4	5	6	7
1	man							
2	innocent							
3	prison							
4	Escape							
5	spider							
6	story							
7	finishes							

A shortcoming of this text is, however, that none of the ideas or themes that are introduced in the opening sentence is carried all the way through to the end. This gives the text a rather abrupt ending.

This opening sentence shows that the writer is capable of constructing a good introduction, but he needed to look more closely at the model text with which he was provided.

In Diagram 4.9, the inverted L is clearly visible. However, the main topic (*man*) is not carried through to the end of the text. What appears to be the main topic (*man*) does not appear in the final two sentences. In fact, none of the reiterated words that appear in the opening sentence appear in the final sentence, meaning that the four shaded squares in the bottom right-hand corner are isolated from the previous sentences. The significance of these observations is now discussed.

4.5.6.3 The central topic or theme

The central topic of this text is *man*, which is carried through to sentence 5 using the repetition of the word *he*. However, this topic is not reiterated in the final part of the text. Unlike in the TBRs, where the presence of reiteration links between the first and final sentences helped to remind the reader of the main ideas of the text, the absence of such links leaves the text hanging in midair, giving the text an abrupt ending.

4.5.6.4 Clusters of repetition

Only one small cluster is noticeable in sentences 4 and 5 in Diagram 4.9, where the writer expands on the story:

4. He is caught and returns to prison.
5. Then he have got an idea when he sees a mole in this time.

However, there is no actual shift in topic: the main topic (*he*, which refers back to *a man*) has been carried over, and *this time* clearly refers to his being returned to prison.

The result of this limited clustering is that there is very little elaboration on precisely what the mole is doing, or of how seeing the mole may give the man the idea to escape. It is left to the reader's knowledge of moles and their tunneling abilities to work out how seeing the mole gives him an idea of escape.

4.5.6.5 Sentences of varied length

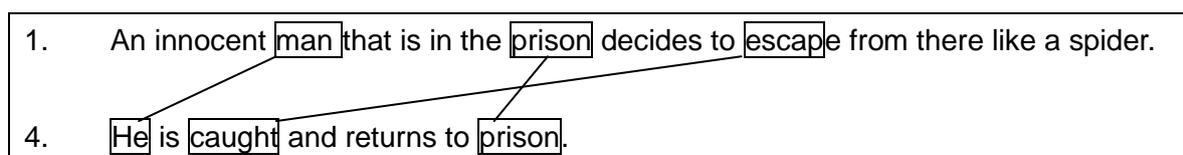
Of the seven sentences in the story, four (sentences 1, 2, 5 and 7) consist of compound sentences. This learner clearly has a good grasp of how to use subordinate clauses effectively. There are also two simple sentences (sentences 3 and 6) which help to emphasize important points, namely, that by escaping and stealing a bicycle he has now become a true thief (sentence 3), and that the story actually ends inconclusively. Sentence 4 is a compound sentence that marks a bridge between his escaping and being back in prison. In this sentence, the writer also successfully elides unnecessary information:

He is caught, and returns to prison.

The elision of *he* demonstrates an understanding of compound sentences, while the use of the passive in the first clause also suggests an ability to use a grammatical structure which is beyond what has actually been taught in class.

As can be seen in Diagram 4.10 below, sentence construction in this text is clearly a strong point in this learner's writing. In fact, sentences 1 and 4 contain three links, which means that these two sentences are bonded.

Diagram 4.10: Bonded sentences 1 and 4 in *Escape*



Sentences 1 and 4 can be read as a brief summary of part of the story. As already mentioned, it is the use of a complex sentence (sentence 1) and a compound sentence (sentence 4), as well as the use of adjectival clauses (*that is in the prison* – sentence 1), which make the bond possible.

Overall, this text is solid in its use of a central topic, as well as in its sentence structure. It also shows hints of capability in all the other criteria, but falls short, firstly, in not containing all the elements of a review (which would have no doubt created more cohesive links), but also in using only a limited variety of cohesive items in the longer cohesive chain (only repetition and reference are used), and not elaborating more in his description of the action.

4.5.6.6 The final sentences

The final two sentences read as follows:

6. The story finishes here.
7. What will happen if the story continues?

Two words (*story* and *continues*) refer to the penultimate sentence. It may also be argued that *story* could be referring to the book as a whole, and thus has a superordinate function. As such, it effectively refers to the beginning, middle and end of the text.

4.5.6.7 Summary

In summary, the text contains a number of effective cohesive chains, and apart from the redundant repetition of the word *he*, the writer is able to vary his use of cohesive devices. However, the cohesive chains are too short and none of them are actually reiterated at the end of the text. It is clear that, as a book review, the text needs to offer the writer's impression of the story. By doing so, it would have created much clearer links to other parts of the text.

4.5.7 ST 4 Phantom of the Opera

Box 4.6 ST 4 *Phantom of the Opera* 114 words: grade – average

1. The name of the book is 'The Phantom of the Opera'.
2. It was nineteenth century at Opera House in Paris.
3. Main characters are the Phantom of the Opera and Christine.
4. The phantom loved her, but she is afraid it.
5. But gradually, she loved him.
6. Finally, they lived far from here together.
7. I enjoyed the book very much, because I like love stories and I like the movie of this book.

4.5.7.1 A variety of cohesive devices used in cohesive chains

The most important chain (see Table 4.30) found in this text is the *Phantom*. This is carried in all but two sentences of the text:

Table 4:30: Lexical chains and cohesive types (1) in *Phantom of the Opera*

Sentence	Topic/idea	Cohesive type
1	Phantom	
3	Phantom	direct repetition
4	Phantom	direct repetition
5	him	reference (pronominal)
6	they (Phantom and Christine)	reference (pronominal)

Two shorter chains (see Tables 4.31 and 4.32) are also found.

Table 4:31: Lexical chains and cohesive types (2) in *Phantom of the Opera*

Sentence	Topic/idea	Cohesive type
3	Christine	
4	her	reference (pronominal)
5	she	reference (pronominal)
6	they (Christine and Phantom)	reference (pronominal)

Table 4:32: Lexical chains and cohesive types (3) in *Phantom of the Opera*

Sentence	Topic/idea	Cohesive type
1	Opera	
2	Opera	direct repetition
3	Opera	direct repetition

While there are lexical chains present, the writer has made use only of direct repetition and reference (pronominal) to create these chains. The writer is able to use some fairly advanced words for this level (*gradually* in sentence 5 and *afraid* in sentence 4). However, he does not yet appear aware of how to use cohesive devices to make the writing more interesting.

4.5.7.2 The opening sentence and repetition links

There are only three words from the opening sentence that are reiterated later in the text. The main topic is *phantom*, while *Christine* and *opera* are secondary topics. Apart from these, *book* is reiterated in the final sentence. However, the opening sentence does not contain any words that may hint at a later shift in topic or references to the middle of the text.

Diagram 4.11: Patterns of reiteration for *Phantom of the Opera*

	Reiterated Type	Sentence Number						
		1	2	3	4	5	6	7
1	Phantom	■		■	■	■	■	
2	Book	■						■
3	Opera	■	■					
4	Paris		■				■	
5	Christine			■	■	■	■	
6	love				■	■		■

As Diagram 4.11 shows, the inverted L that was visible in all the previous texts analyzed is not as clear in this diagram. There appear to be two main topics, but neither of these begin

in the first sentence or end in the final sentence. There is also a break in what appears to be the main topic (*phantom*) in sentence 2.

4.5.7.3 The central topic or theme

As was noted in §4.5.7.1 above, the main theme of the text is the *phantom*; Diagram 4.11 above shows the patterns of reiteration in the text.

The theme of the *phantom* is reiterated in sentences 1, 3, 4, 5 and 6. The second most important theme is that of *Christine*, which appears in sentences 3, 4, 5 and 6. The main problem with this text, however, does not concern the main theme, but rather a lack of secondary topics. This is examined in the next section.

4.5.7.4 Clusters of repetition

Clusters of repetition are very limited. The only cluster that can be seen occurs around sentences 1, 2 and 3 (see Table 4.33).

Table 4:33: Repetition clusters in *Phantom of the Opera*

Sentence	Idea 1	Idea 2	Idea 3	Idea 4	Idea 5
1	Phantom	Opera			
2		Opera		Paris	
3	Phantom	Opera			Christine

In this cluster, the main theme is in fact the opera. This cluster reveals where the action takes place and introduces the other main character, *Christine*. As such, it is very effective in not only anchoring the known information in the word *opera*, but also introducing new information.

This is the only cluster that occurs; it suggests that while the writer is able to elaborate and explain, he needs to apply this to other parts of the text.

4.5.7.5 Sentences of varied length

The text contains mostly simple sentences and very few complex or compound sentences. Of the seven sentences in the text, five (sentences 1, 2, 3, 5 and 6) are simple. However, the writer uses an adjectival and an adverbial phrase in two of his sentences, which help to add important information. The boxed areas indicate these adjectival (sentence 2) and adverbial (sentence 6) phrases:

Sentence

2. It was nineteenth century at Opera House in Paris.
6. Finally, they lived far from here together.

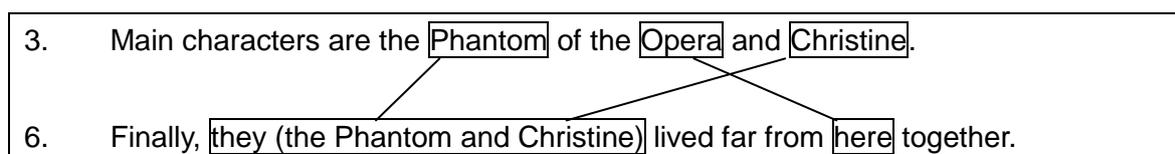
This goes some way to mitigating the overuse of simple sentences. The writer also makes good use of adverbs, which had only been studied relatively shortly before writing this essay.

The only compound sentence in the text is sentence 4, while the only complex sentence is the final one, sentence 7:

I enjoyed the book very much, because I like love stories and I like the movie of this book

As Diagram 4.12 shows, no bonds were found in this text, although it could be argued that sentences 3 and 6 are bonded, if one accepts that *they* in sentence 6 refers both to the *phantom* and to *Christine*:

Diagram 4.12: Linked sentences 3 and 6 in *Phantom of the Opera*



Since narrative texts make use of time markers (such as *finally* in the case of Diagram 4.12 above, the presence of bonds also indicates the central themes of the text.

4.5.7.6 The final sentence

The final sentence contains three links (*book*, *love* and *stories*).

The final sentence reads as follows:

7. I enjoyed the book very much, because I like love stories and I like the movie of this book.

The final sentence contains the word *book* (twice) which links back to *book* in sentence 1. Moreover, the word *love* is a repetition of *love* in sentence 4. Finally, *stories* refers to the story as a whole. In addition to performing an important role in offering the reader's impression of the story, the final sentence in this text also provides links to the middle and end of the text. This is a strong sentence, because it elaborates on the main clause, namely, *I enjoyed this book*. Firstly, it provides an adverbial phrase (*very much*) to describe the degree to which he liked the book. It then offers two reasons in a compound adverbial clause of reason. Many learners at this level have a tendency to begin a new sentence with *because*, creating a subordinate clause fragment, but not his writer.

4.5.7.7 Summary

Overall, this text has a number of strong points. It contains a main topic, forms cohesive chains with *phantom* and *Christine*, and ends with a strong final sentence. The writer also demonstrates an ability to use adverbs correctly. The final sentence indicates that the writer is capable of writing a complex sentence. These points are the reason why the writer was given an average rather than a poor grade. However, the story lacks longer, more complex sentences. This no doubt accounts for the lack of elaboration and consequently a very

limited cohesive cluster with no real move away from the main topic. It also employs a limited variety of cohesive devices.

4.5.8 ST 5 The Scarf

Box 4.7 ST 5 *The Scarf* 63 words: grade – below average

1. I read the book which name is "The Scarf"
2. This book's main character is Anneena.
3. She made a very long scarf in winter.
4. Anneena and Anneena's friends went to the park.
5. There was a frozen pond in the park.
6. A boy falled in the frozen pond.
7. Anneena threw her long scarf to boy to help him.
8. The boy was safe.
9. Anneena was glad.

4.5.8.1 A variety of cohesive devices used in cohesive chains

This essay contains just two short chains (see Tables 4.34 and 4.35 below) in which reference and direct repetition are used.

Table 4:34: Lexical chains and cohesive types (1) in *The Scarf*

Sentence	Topic/idea	Cohesive type
2	Anneena	
3	she	reference (pronominal)
4	Anneena	direct repetition

Table 4:35: Lexical chains and cohesive types (2) in *The Scarf*

Sentence	Topic/idea	Cohesive type
6	boy	
7	boy	direct repetition
8	boy	direct repetition

Other repeated words are:

- book (sentences 1 and 2)
- scarf (sentences 1 and 7)
- long (sentences 1 and 7)
- park (sentences 4 and 5)
- frozen (sentences 5 and 6)
- pond (sentences 5 and 6)

Apart from reference and repetition, no other cohesive devices are used. This means that there is little or no variation in the use of cohesive devices.

4.5.8.2 The opening sentence and repetition links

I read the book which name is *The Scarf*

While the first sentence is adequate in terms of suggesting that the learner is able to use the recently learned subordinate clause, it becomes clear from the outset that this writer can only provide a little new information in each sentence. The reason for this becomes clear as one examines the other criteria below. Furthermore, apart from *book* and *scarf*, there are no other words that link to other parts of the text.

Diagram 4.13: Patterns of reiteration for *The Scarf*

	Reiterated Type	Sentence Number								
		1	2	3	4	5	6	7	8	9
1	book									
2	scarf									
3	long									
4	Anneena									
5	park									
6	frozen									
7	pond									
8	boy									

As can be seen in Diagram 4.13 above, the inverted L pattern does not occur in this text, mainly because the writer simply gives the name of the book in the opening sentence. Only two words are reiterated in the text, namely, *book* (in sentence 2) and *scarf* (sentences 3 and 7). This, however, is mitigated by sentence 2, which, as mentioned, is not only linked to sentence 1 through the word *book*, but also contains the main topic, *Anneena*, and the word *long* – later repeated in sentence 7.

4.5.8.3 The central topic or theme

Although a main topic or theme exists, its presence is somewhat limited. The main topic of this text is *Anneena*, which is reiterated in sentences 3, 4, 7 and 9, while *scarf* (sentences 1, 3 and 7) may be seen as a secondary theme. Unlike other the more successful texts, the main topic does not appear in the first sentence; although the opening sentence introduces the theme of *scarf*, this is only reiterated in sentences 3 and 7.

4.5.8.4 Clusters of repetition

There are no clusters of repetition in this text. The closest this text comes to forming any cluster is from sentences 4 to 7. This is laid out in Table 4.36 below:

4. Anneena and Anneena's friends went to the park.
5. There was a frozen pond in the park.
6. A boy falled in the frozen pond.
7. Anneena threw her long scarf to boy to help him.

Table 4:36: Near-repetition cluster in *The Scarf*

Sentence	Idea 1	Idea 2	Idea 3	Idea 4	Idea 5	Idea 6	Idea 7
4	Anneena	friends	park				
5				park	frozen	pond	
6					frozen	pond	boy
7	Anneena					boy	

Sentence 4 contains the main topic (*Anneena*), as well as *park*. The theme of the *park* is then carried through to sentence 5, where *frozen* and *pond* are also added. These two words are then repeated in sentence 6, where *boy* is added. Sentence 7 then repeats the theme of the *boy*, as well as reintroducing the main topic, *Anneena*. As can be seen, each sentence contains a link with the next, and also adds a little more information. This indicates that this writer is able to set out a narrative, but only through direct repetition. Moreover, the move away from the main topic is only two sentences long, and only introduces *frozen pond* and *the boy*.

4.5.8.5 Sentences of varied length

All sentences, barring 1 and 7, are simple sentences. Moreover, as was noted in §4.5.8.5 above, even sentence 1, which contains a relative clause, does not provide very much new information. While the sentences are generally clear, there is a great deal of redundant repetition. The only exception is sentence 7 which demonstrates the importance of more complex sentences by containing four links to other sentences.

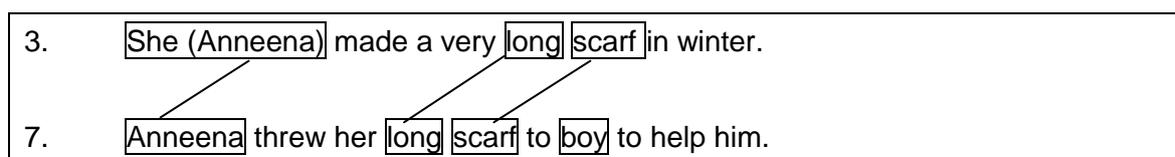
7. Anneena threw her long scarf to boy to help him.

The boxed words are linked to the following sentences:

Anneena	sentences 2, 3, 4, 8
long	sentence 3
scarf	sentence 3
boy	sentence 6

Because sentence 7 contains three words which also appear in sentence 3, the two sentences can be said to form a bond.

Diagram 4.14: Bonded sentences 3 and 7 in *The Scarf*



Yet again, as Diagram 4.14 shows, sentences 3 and 7 make sense (albeit without much detail added to the narrative) when juxtaposed. Sentence 7 contains a subordinate clause (*to help him*) which helps to explain why she threw the scarf, and also creates a link between *boy* in sentence 7 to *boy* in sentence 6. This suggests that the learner was capable of creating more complex sentences, but had not perfected or mastered this construction. At this stage, he was still relying heavily on simple sentences.

4.5.8.6 The final sentence

The final sentence does not link back to the beginning, middle or ending of the text. It contains only one link (*Anneena*) to the previous sentences. The same is true of the penultimate sentence (*The boy was safe*), which is linked to sentence 7. This means that the text ends with the summary of the narrative, and does not offer any impression of the text. The result is a text that ends somewhat in midair, without any final reiteration of the main points.

4.5.8.7 Summary

The writer of this text displayed some ability to relate the story of the book that he read. The low grade given for this text was mainly due to the fact that it felt somewhat disjointed and choppy to read. The reason for this disjointedness lies in the very limited variety of cohesive devices and lack of significant lexical chains. The opening sentence contains no words that are reiterated elsewhere in the text. While there is a central topic, it is not sustained consistently throughout the text. There are no discernible clusters of reiteration,

although the writer did attempt to elaborate around sentences 4 and 7 by using the adverbial phrases *to the park* in sentence 4 and *to help him* in sentence 7:

4. Anneena and Anneena's friends went to the park.
7. Anneena threw her long scarf to boy to help him.

The main failing of this text is its inadequate use of complex and compound sentences, of which only two could be found. Simple sentences by their nature can only offer a limited amount of information at a time. However, the one complex and one compound sentence that do occur suggest that the learner may be able to construct them, and will perhaps practice using a greater number in future essays. Finally, the last sentence, which did not offer the writer's impression of the story, contained only two links to different parts of the text, making the text seem incomplete.

4.5.9 ST 6 Sangokushi

Box 4.8 Text for ST 6 *Sangokushi* 64 words: grade – below average

1. Name of book is "Sangokushi".
2. It happen very old China.
3. This book's main character is Ryubi and Kanu and Chouhi.
4. They was leaving Syoku and war between Go and Gi.
5. Gi is strongest of the three, because it had many people.
6. I like best Syoku of the three.
7. To I like Koumei.
8. He has best brain and he win to the Gi.
9. Everyone should read it.

4.5.9.1 A variety of cohesive devices used in cohesive chains

No words are repeated in more than two successive sentences, as can be seen in Table 4.37:

Table 4:37: Lexical chains and cohesive types for *Sangokushi*

Sentence	Topic/idea	Cohesive type
1	Sangokushi	
2.	It	reference (pronominal)
3.	Ryubi, Kanu, Chouhi	
4	They	reference (pronominal)
4.	Gi	
5	Gi	direct repetition
6.	the three	
7.	the three	direct repetition

The use of *the three* was marked as a superordinate of Ryubi, Kanu and Chouhi in the quantitative analysis.

It is clear, however, that the writer relies mainly on repetition and reference to achieve cohesion.

4.5.9.2 The opening sentence and repetition links

1. Name of book is "Sangokushi".

The first sentence contains two words that are linked to other sentences, namely, *book* and *Sangokushi*.

book reiterated in sentences 3 and 9
Sangokushi reiterated in sentence 2

These are the only two words that are reiterated in subsequent sentences. The idea of *book* does not elaborate on the story in any way. Since the idea of *Sangokushi* is repeated only in the second sentence, no other ideas concerning the story itself are continued in the middle or the end of the text.

Diagram 4.15: Patterns of reiteration for *Sangokushi*

	Reiterated Type	Sentence Number								
		1	2	3	4	5	6	7	8	9
1	book									
2	Sangokushi									
3	Ryubi									
4	Kanu				X					
5	Chouhi									
6	Shyoku									
7	Go					X				
8	Gi									
9	Kyomei									

Note: The boxes marked with an X indicate a word that refers to all of the three shaded boxes in the previous sentence.

As can be seen on Diagram 4.15 above, the inverted L pattern, which was so prominent in the stronger texts, is absent. Firstly, only two words from the first sentence are reiterated in the text and, secondly, neither the opening sentence nor the second sentence introduces a central topic or theme of the text. In the second sentence, we are told that the action takes place in China. Although sentence 3 repeats the word *book* from sentence 1, this is the only tenuous link to the previous sentences.

4.5.9.3 The central topic or theme

As has already been mentioned in §4.5.9.1, there is no clear topic that is developed through the text. Moreover, as already noted, no words are reiterated more than once in successive sentences. It is therefore difficult or almost impossible to get an idea of the main idea or topic of the text.

4.5.9.4 Clusters of repetition

There appears to be something of a cluster in sentences 4, 5 and 6. However, this is due largely to a great deal of repetition, without very much clear elaboration of who precisely characters are, or the nature of the places that have been mentioned. Furthermore, unlike

the stronger texts, this text is not anchored in a main topic from which it can shift when it needs to elaborate, or to which it can return. This is clear in Table 4.38:

Table 4:38: Near-repetition clustering in *Sangokushi*

Sentence	Idea 1	Idea 2	Idea 3	Idea 4	Idea 4	Idea 6	Idea 7	Idea 8	Idea 9
3	Ryubi	Kano	Chouhi						
4	They			Syoku		Go	Gi		
5					the three			Gi	
6				Syok u					
7									Koumei

While the more successful texts clearly created a main topic by linking their sentences using cohesive devices, or included a sentence with references to both the main and the new topic (see the ST1 *Tom Sawyer* and ST2 *The Jumble Sale*), this text moves away from one topic, but then fails to return to it. Instead, it simply continues to add new information, such as *Syoku*, *Go*, *Gi* in sentence 4, or *Koumei* in sentence 7, without explaining to the reader who or what these are.

4.5.9.5 Sentences of varied length

Of the nine sentences in the text, only one is a complex sentence (sentence 5) and one a compound (sentence 8). All the others are simple sentences. Moreover, as has already been noted in 4.5.9.2 the result of this is that the information is presented haphazardly with very little reference to what came before. For example:

6. I like best Syoku of the three.
7. To I like Koumei.
8. He has best brain and he win to the Gi.

The reason for beginning the penultimate sentence with 'To' is that the students are introduced to the phrase 'in order to' (which is then shortened to 'to') for expression of purpose before they are introduced to 'because', which explains pre-existing reasons. This leads to many students confusing the use of 'to' and 'because'. Many tend to think that one can use 'to' to explain both purpose and pre-existing reasons. This is also possibly due to language interference, since there is a Japanese word (-*no tame*) which covers both *to* and *because*. Despite my best efforts, some of the weaker learners still tend to make this error.

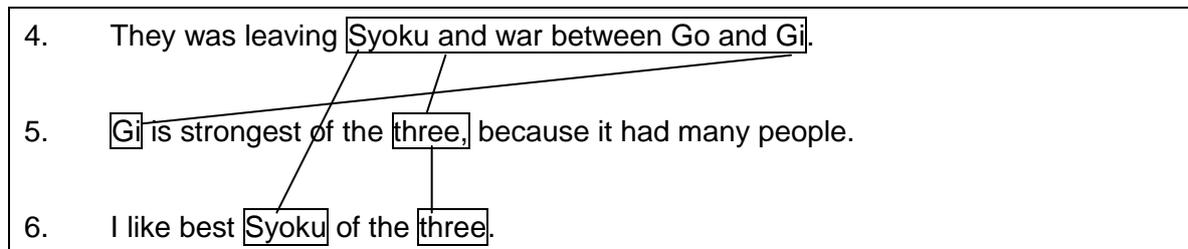
However, the writer of ST 6 uses *because* appropriately in sentence 5:

5. Gi is strongest of the three, because it had many people.

This suggests that he may be able to use the structure in some cases but not in others.

As can be seen in Diagram 4.16 below, there are very few links between sentences in this text:

Diagram 4.16: Sentence links 4, 5 and 6 in *Sangokushi*



This is about the closest the text comes to forming some kind of bond between sentences.

4.5.9.6 The final sentence

The final sentence reads:

9. Everyone should read it.

The *it* in the final sentence refers back to *book* in sentences 3 and 1. However, it could also be interpreted as referring to Koumei's *brain* instead. This is the only link to any other parts

of the text, which means that the writer has failed to summarize the text, or remind the reader of the key ideas it contains. It was noted that some of the STs did in fact summarize in the penultimate sentences (*Tom Sawyer* and *The Jumble Sale*). However, the penultimate sentence in this text (*He has best brain and he win to the Gi*) also contains very limited links (*Kyomei* to sentences 6 and 7; and *Gi* to sentences 4 and 6). Moreover, the use of *it* in this sentence is strange, because reference is usually used in adjacent sentences, or when it is clear precisely to whom the writer is referring. In this case, the reference item is separated by five sentences. Finally, the writer does not say why he thinks 'everyone should read it'. A simple explanation would go a long way to reiterating some of the most important points of the story, and further creating links to other parts of the text.

4.5.9.7 Summary

In summary, this text has all the components of a book review. It offers a setting, mentions the main characters, shows the possible problem, and gives the writer's impression of the text. The writer also attempted to use grammatical structures that had recently been studied, with varying degrees of success.

However, the text contains none of the features of successful writing identified in §4.4. The result is that it is often unclear precisely to what the items are referring. It would appear that this writer is assuming too much knowledge on the part of the reader. As was noted in Chapter 1 (§1.2.3), attention to audience is an important part of composition, the writer needs to consider who might be reading the text and decide how much or how little information needs to be provided. The writer of this text is expecting the reader to be able to connect *Ryubi* and *Kanu* and *Chouhi* with *Syoku* (sentence 3), *Go* (sentence 4) and *Gi* (sentence 4). While he does use *they* in sentences 4 and 5, it is unclear precisely to whom this refers.

4.5.9.8 Summary of qualitative analysis

Table 4.39 below provides a summary of how the six STs compare to the TBRs.

Table 4.39: A summary of how the six STs compare to the TBRs

Features of text	Text Number					
	1	2	3	4	5	6
Cohesive Chains and variety of cohesive devices	Yes	Yes	Few	Few	Few	No
Opening sentence use of reiterated ideas	Yes	Yes	Few	Few	No	No
Central topic or theme	Yes	Yes	Yes	Yes	No	Few
Clusters of reiteration	Yes	Yes	Few	Few	No	Few
Varying sentence length	Yes	Yes	Yes	Few	Few	No
Final sentence links to other parts of the text	Few	Few	Few	Yes	No	No

It is clear that the stronger texts contain all or most of the features identified in the TBRs, the average texts contain some or some to a limited extent, while the weaker texts contain few, if any, of these features. If we regard TBRs as sound models, this demonstrates that the features identified are indeed important in establishing cohesion. The findings suggest that learners at this level are capable of including these features in their writing. Finally, they provide an indication of why the weaker texts were graded lower, and point the way to how these problems might be remedied.

4.6 FINDINGS RELATED TO RESEARCH QUESTIONS 5 AND QUESTIONS 6:

Research Question 5 relates to cohesive errors and sought to ascertain which cohesive devices in the STs are used inappropriately or inaccurately. Question 6 examined the reasons for these errors.

Table 4.40: Total number of cohesive errors in the STs (excluding repetition) in order of the number of errors

Text no	Text Title	Conjunction	Exophoric Ref	Unclear Ref	Total
1	Phantom of the Opera			1	1
6	Sangokushi	1			1
7	My Home		1		1
9	Alice in Wonderland	1	2		3
10	The Lost Key	5	1		6
12	The Red Planet	2	1		3
14	A Rabbit and a Turtle	2	1		3
15	Soup with Obby			2	2
18	The Jumble Sale	1		1	2
21	Pocket Money	2	1		3
22	The Surprise		2		2
23	Biscuit			4	4
24	The Rubbish Monster	2			2
26	Jojo and the Football		1		1
27	Vanishing Cream		2		2
29	Wind and Fire	1			1
33	Camping Adventure	1			1
34	The Red Planet	1			1
35	Spock and Donkey	1			1
36	The Cold Day		1		1
	Total	20	13	8	41

In all, 41 errors of cohesion were found in 20 out of the total of 36 STs. This total includes 20 conjunctions that were also included in the main count, despite being used erroneously. As was noted in §4.3.1.1, the reason for including them in the main count was to ascertain whether conjunction was being used, irrespective of whether it was used correctly or not. However the above total of 41 errors excludes redundant repetition, which falls under repetition, and is difficult to quantify. It should also be noted that many of these errors

appeared in texts other than the six analyzed in the qualitative analysis. Therefore, texts where the cohesive errors were found will be included in this analysis. Table 4.40 shows the number of errors in each text, including conjunction errors.

These results indicate that half of the STs contained at least one error, while a relatively small number of texts accounted for all the errors committed. Twenty texts or 56% of the sample contained at least one error. Fourteen texts contained two or more cohesion errors, and eight texts contained more than two. The eight texts containing three or more errors of cohesion accounted for 37 of the 55 errors committed. This means that only eight learners or 22% of the sample accounted for 67% of the cohesion errors, excluding redundant repetition. Thus errors were concentrated in a few texts, rather than being evenly distributed.

What is important to note is that this section of the study is not focused on the quality of the texts per se, but rather on the errors and problems in the use of cohesive devices. In some cases the success of the text was affected by these errors, such as in the unclear referencing found in *Biscuit* (see Box 4.9 below). The lack of clear referencing makes it difficult to understand who is doing what. However in other cases, such as in *The Jumble Sale* (§4.5.5), they had no major effect on an otherwise well-constructed text.

The types of error or deviation identified can be divided into the following four categories: (i) exophoric reference; (ii) unclear referencing; (iii) conjunctions; and (iv) redundant (as opposed to effective) repetition. This section will examine each of these in turn. All errors with the exception of repetition can be accounted for quantitatively. This means that it is possible to count the number of errors. With repetition, on the other hand, redundancy is more of a qualitative issue because it is not always the repetition itself that is the problem, but where it occurs in the sentence. In this section, repetition will be addressed by

examining actual samples of text and showing why the repetition was found to be redundant.

4.6.1 Exophoric reference

In certain cases, exophoric reference (namely, reference to someone or something outside the text) may well be appropriate. This is the case in texts or fragments of text where it is safe to assume that the reader is familiar with the referent. For example, in the textbook, the opening sentence of the *Pocahontas* story reads as follows:

One of the colonists in Jamestown was Captain John Smith. (Text E)

If this text is read in isolation, it may be difficult to work out who *the colonists* refers to. However, in this case, *the colonists* is in fact referring to the previous reading on Christopher Columbus and the colonization of America. It can be argued that the entire textbook itself is one text, which would mean that the reference is not entirely exophoric, but in fact helps to maintain cohesion throughout the textbook as a whole.

A quarter of the sample of 36 texts used exophoric referencing inappropriately at least once, suggesting that this was a common area of difficulty among learners at this level. Among the nine texts that contained exophoric referencing, four texts contained two errors. What follows is a description of how the errors affect the texts.

In ST 14 (*The Rabbit and the Turtle*), the first sentence begins as follows:

The name of the book is *a Rabbit and a Turtle*.

In the above sentence, *the book* may refer to *the book which I read for this assignment*. However, anyone unfamiliar with this information would find it difficult to understand precisely which book was being referred to.

In ST 19 (*Escape!*), the text opens with:

The main character is Brown.

Again, it is impossible for a reader who is not familiar with the nature of the assignment to understand the context in which the main character exists. In this case, the writer appears to have simply taken one of the sentences from the model text (see Appendix 2) and modified it for his essay.

Another example of exophoric reference can be seen in Text 21 (*Pocket Money*), and in Text 22 (*The Surprise*):

Dad maked a jobs chart. Dad say 'You can make extra pocket money'. (ST21)
Ben and Ben's family do something (ST 22)

In both cases, the characters are introduced without giving any hint as to who they are or how they fit into the story. In ST 21, *Dad* is introduced quite abruptly. While one would no doubt assume that he is the father of the children, it would be more appropriate to introduce him as such. ST 22 contains an exophoric reference in the form of *Ben*, without first giving some indication of who Ben is.

These errors suggest that some learners are not able to use the model text effectively to introduce their stories and to provide context for their characters. In the examples cited above, the learners have simply answered questions from the rubric that I provided, or have taken extracts of the model text without properly understanding how they may be used. One possible reason for this is that they are not as yet used to writing essays, and do not have the confidence to create more complex sentences.

4.6.2 Vague or ambiguous referencing

In all there were eight cases of unclear referencing in four STs. Half of the cases of unclear referencing were to be found in only one text. Thus, while 90% of the learners are able to use referencing properly, at least one learner (the writer of ST 23) had some serious difficulties in this regard. This learner seemed to understand how to write grammatical sentences, but had problems structuring his essay. The lack of referencing in ST 23 is discussed in more detail below. Also of interest is that referencing problems occur in the stronger texts as well: ST 18 (*The Jumble Sale*), which was identified in the previous section (§4.5.5) as a highly successful text contains one unclear reference, as does the partially successful ST 1 (*Phantom of the Opera*, §4.5.7)

Since Text 23 contains half of the unclear referencing cases identified, it warrants special attention. The full text appears in Box 4.9 below.

Box 4.9 Full text of ST 23 - *Biscuit* 74 words

1. The dog and her horse.
2. "Biscuit time for bed" she said.
3. But he want to play.
4. He go to by her.
5. He said "I want to snack".
6. "I want to drink hear a story, my blanket ..."
7. This book name is Biscuit.
8. Biscuit is the main characters.
9. He is the dog.
10. This book has six series.
11. It is short story one by one.
12. It is new kind story, but there are a lot of new sentence.

The first sentence is already problematic because it does not contain a verb. It is also unclear whether *her* refers to the dog or is perhaps an exophoric reference to a girl or woman outside the text. Similarly, sentence 2 contains a reference to *her* in sentence 1,

she. Although *she* in sentence 2 appears to be referring to the *her* in sentence 1, it could also be referring to the dog or the horse. Again, in sentence 3, it is unclear whether *he* is referring to the dog or the horse mentioned in sentence 1. Finally, the *I* in sentence 6 is a reiteration of the *I* in sentence 5. However, it is unclear up to this point precisely whether it is the dog or the horse to which *he* and *I* are referring. This text illustrates the importance of initially naming the characters. In this case, the person referred to as *she* should have been given a name or some kind of description.

This text also highlights the importance of getting the structure of text right, in other words understanding the different phases (introduction, body, conclusion) in a text. In this text there is no clear introduction. The writer simply launches into the story. Only at sentences 7 to 9 is it explained who the main characters are. Had these sentences occurred earlier, some of the confusion could have been avoided.

While texts revealed that writers were generally able to use pronouns, there were some cases in which it was not entirely clear to whom or to what the pronoun referred. For example, in Text 21 (*Pocket Money*) it is difficult to ascertain whether the *she* in sentence 7 refers to *Biff* or *a model*.

11. Some time they find a model at the sea.
12. Model take a photo.
13. But model forgot a camera.
14. Biff bring camera.
15. Model give moneys.
16. But she didn't take moneys.

Although sentences 15 and 16 are linked by the word *money*, it is still difficult to ascertain at first glance whether the *she* in sentence 16 is referring to *model* in sentence 15, or to *Biff* in sentence 14. One of the reasons why there is an unclear reference in this text is that the writer has not included a direct object (*her*) in sentence 15. This would have created a link

to *Biff* in sentence 14. It would also then be clearer in sentence 16 that *she* was referring to *Biff* rather than *model*.

4.6.3 Conjunction

As noted in the quantitative analysis (§4.3.1.1), the use of conjunctions in the STs was quite frequent but limited to a very narrow range of types. Moreover, there were many cases where sentences began with conjunctions that were not appropriate. The TBRs on the other hand tended to use conjunctions more to connect clauses within sentences, rather than as a way of referring back to previous sentences. They also used a far wider variety. It was noted in the quantitative findings (§4.3.1.1) that 56 conjunctions were identified in the STs, of which 20 were used inappropriately. The number of errors in conjunction per ST can be seen in Table 4.40 (§4.6).

There are two possible reasons for these errors:

1. Learners are modeling their sentences on those found in the textbook without actually understanding when these can be used appropriately.
2. Learners may understand certain conjunctions found in the textbook but these have yet to become part of their active vocabulary.

These reasons are examined below, using examples from the STs.

4.6.3.1 Model sentences from the textbook

The textbook itself may be a reason that many of the weaker compositions comprised mostly simple sentences, often a source of redundancy. Many STs started free-standing subordinate clauses with *because*, probably imitating dialogues in the textbook. In spoken English, people readily speak in subordinate clauses. For example, in a model dialogue in the textbook (*Progress 21 Book 2*), the following exchange is found:

Grandma: Really? How long is he going to stay?
Tom: Until the beginning of July. (Flynn 2011a:6)

In Lesson 2 Scene 3

Dr. Green: If you don't finish your vegetables, you can't have any dessert.
Tom: But we came here because the desserts are so great. (Flynn 2011a:20)

Moreover, the answers to the reading sections in the teacher's manual also start sentences with *because*:

Why was (Christopher Columbus) able to (sail west)?
Because the Queen of Spain believed him and gave him three ships (Flynn 2011b:15).

These sentences are not necessarily incorrect in the context of a conversation, or where the need to present the answers to a comprehension question is more pressing than the need to write complete sentences. However, confusion arises when it comes to free writing exercises where the use of free-standing subordinate clauses is not generally considered good writing practice, particularly in formal writing.

Four of the 20 errors in conjunction involved the use of *because* in a free-standing subordinate clause:

I enjoyed this book. Because the magic key's power is very interesting. (ST 10)

As this example demonstrates, it is clear that the learner thinks that, since free-standing subordinate clauses occur in the textbook, it is possible to use them in written compositions. Indeed, this has been something of a perennial problem when teaching despite my best efforts.

In ST 10, a number of sentences begin with *and*:

The key showed Biff and his friends many children's dreams. *And* the key was shining ... One day Biff shopping. He was playing in the park on the way to the shop. *And* the key fell out from pocket. Biff and his friends were looking for the key. *And* they found it.

Again, part of the reason for the overuse of *and* in this composition is due to the fact that there are sentences in the textbook that begin with *and*. The main reason is that learners simply did not, at this point in their learning, have any other words at their disposal with which to connect the sentences. The writer of ST 10 seems to understand that a conjunction may be used but only has the basic 'and' at his disposal.

4.6.3.2 Limited active knowledge of conjunctions

The fact that learners had only started using conjunctions a few weeks before the essays were written may also account for the overabundance of conjunctions such as *and*, *so* and *but*. At this point, learners had been exposed to a number of conjunctions in the textbook, but for many these were yet to become part of their active vocabulary. This is clearly evident in the results.

In some cases, the learners were clearly attempting to expand their use of conjunctions, but were making errors in vocabulary in the process. Indeed, 13 different conjunctions were used, of which four were judged to be vocabulary errors:

after instead of *later*: 2 cases

and in place of *then*: 4 cases

starting a sentence with *And*: 7 cases

Sometime instead of *later*: 1 case

A fifth sentence was also a free-standing subordinate clause that used *to* instead of *because* to indicate a prior reason.

I like best Syoku of the three. To I like Koumei. (Text 6)

As mentioned above (§4.5.9.5), this error occurred as a result of the learner confusing the phrase *in order to* with *because*.

Another possible reason for some of these errors may be that, having learned conjunctions relatively recently, learners tended towards overuse, something which often happens when learners are introduced to a new grammatical structure. These errors indicate that the learners had some understanding of conjunctions but still required practice in using them more sparingly and effectively.

4.6.4 Redundant repetition

Since redundant repetition is difficult to quantify, it was not included in the error count. However, the qualitative analysis of STs found a number of cases of redundancy.

In ST 5 (*The Scarf*), the writer relies too heavily on the main character's name (Anneena):

1. This book's main character is Anneena.
2. She made a very long scarf in winter.
3. Anneena and Anneena's friends went to the park.

The writer is clearly unable or reluctant to use possessive pronouns. These would have been taught the previous year in J1, so it is surprising that he still appears to be struggling in this regard. The writer has avoided using pronouns, as well as compound sentences.

A similar situation can be seen in ST 20 (*Greg's Microscope*). Here, the writer has avoided using personal pronouns. While dealing with the narrative part of the text, he repeats the name *Greg*, but in the part where he is expected to offer an opinion, he combines sentences, and makes better use of personal pronouns.

The text is presented in its entirety below:

1. Main characters are Greg and his parents and his friends.
2. Greg wanted a microscope because he watch many kinds of materials.
3. Greg's parents gave him a microphone*.
4. Greg was happy.
5. Greg used to look many kinds of materials, and surprised at them.
6. This story is peaceful, so you can read it easy, but it's slight long.
7. I would enjoy this book, because Greg is still young child.
8. He is interesting and cute.
9. I recommend this book.

*The student wrote *microphone*, although he clearly meant *microscope*

While the writing is not grammatically incorrect, it becomes repetitive and does not reflect the natural rhythm of native English. Four consecutive sentences (2–5) begin with *Greg*. In fact, Greg's name appears in seven of the 10 sentences, rather than using pronominals to relieve the monotony.

It is important to note that the introduction of a particular structure to a group of learners does not mean that all students will acquire it at the same time. There is a considerable variation in the pace at which learners acquire a language. This can be due to numerous factors, such as a natural talent for language, motivation in the language, as well as one's attitude to language learning in general. All of these may explain why some learners are able to create fairly sophisticated texts, while others are still struggling with basic grammatical structures and simple sentences.

4.7 SUMMARY AND CONCLUSION

This section summarizes the main points of the study by addressing the five main research questions.

The findings showed that there was a significantly higher frequency of ellipsis and synonyms in the TBRs than in the STs, while there was a significantly higher frequency of conjunction and reference in the STs relative to the TBRs. In all other devices, the differences in frequency between the TBRs and the STs were not significant. A deeper investigation into the actual words used found that the TBRs tended to use more types and reflected a richer vocabulary, whereas the STs were noticeably limited in the number of types used. This suggests that the writers were not able to apply all the vocabulary that had been presented in the textbook up to that point in the curriculum.

Given that the TBRs were essentially model texts, the question then arose of the extent to which the features of reiteration they contained were to be found in the STs. The qualitative analysis revealed how the more successful STs contained most or all of the same features and patterns of reiteration as the TBRs. It also showed that moderately successful texts showed some, but not all of these features, while the weaker texts contained few, if any.

This study identified four main errors in cohesion in the STs, namely, exophoric reference, redundant repetition, ambiguous referencing and conjunction. It was noted that more than half of the texts contained at least one cohesion error, but that most of the errors were concentrated in the weaker texts. While many of these errors did not necessarily affect the overall quality or cohesion of the texts they merited examination because of what they revealed about how the learners went about constructing texts. They showed that some learners clearly needed more help and practice in understanding how cohesive devices could help to improve texts. While almost all the learners needed help with one or two problems, a few learners needed help with all the errors cited. This analysis therefore helped not only to isolate the errors but also to identify which learners were making them.

The following chapter discusses the contributions of the study, the implications of its findings and the limitations of the research, and makes recommendations and suggestions for further research in the future.

Chapter 5: Conclusion

5.1 INTRODUCTION

The chapter starts by providing a brief summary of the aims and findings of the study. It then discusses the findings, showing how they support previous research, as well as accounting for some of the possible reasons for these results. Following this discussion, it looks at some ways in which the study has contributed to furthering our understanding of cohesion in the writing of Japanese junior high school learners. Next, it discusses the limitations of the study, after which it makes suggestions as to how the findings might be practically applied. Finally, recommendations are made for further research.

5.2 SUMMARY AND DISCUSSION OF THE FINDINGS

The main thesis of this study, as presented in Chapter 1 (§1.8) was as follows:

- It is possible to gain a clearer picture of the strengths and weaknesses of Japanese Grade 2 junior high school learners' compositions by comparing the cohesive devices found in them to those found in the reading sections of the textbook they are using.

This section summarizes the findings with respect to the research questions discussed in Chapter 1 (§1.7)

5.2.1 Findings for Research Question 1

The findings discussed here relate to Research Question 1: What cohesive devices are found in the textbook readings (TBRs), and what cohesive devices are found in the student texts (STs)?

As noted in Chapter 4 (§4.2), the raw data showed that the frequencies of cohesive devices as a proportion of each corpus were quite similar. This suggested that, at face value at least,

there were no major differences between the two texts. However, these initial results still did not account for the obvious differences between the TBRs and the STs in the quality of the texts. In order to determine this, some statistical analysis and qualitative analysis were necessary.

5.2.2 Findings for Research Question 2

This section includes a discussion on the findings relating to Research Question 2: Is there a difference in frequency in the use of cohesive devices in the TBRs and in STs?

No significant difference was found between the TBRs and the STs in the total frequency of cohesive devices. This supports similar findings by Abadiano (1995), Hinkel (2001), Castro (2004), Mojica (2006), Chen (2007), Nakao (2009) and Okuda (2012).

With respect to individual cohesive devices, there was a significantly higher frequency of ellipsis and synonyms in the TBRs, and a significantly higher frequency of conjunction and reference in the STs. No significant difference was found between the TBRs and STs in the use of substitution, superordinates, collocation, repetition and general items. Since there were very few cases of substitution and general items, these were disregarded, even though they are important devices to be used in more advanced texts. These results are now discussed in turn.

5.2.2.1 Ellipsis

Since ellipsis is often found in dialog, and the TBRs contained a high percentage of dialog (about 43%), the significantly higher frequency of ellipsis in the TBRs was not unexpected. It is tempting, as Nakao (2009) did, to disregard ellipsis entirely. However, despite this bias, ellipsis remains an important device to teach because it helps to free the text of redundant information

5.2.2.2 Synonyms

The significantly higher frequency of synonyms in the TBRs supports the findings of Okuda (2012), in which the higher rated compositions contained significantly more of this device.

In the case of synonyms, the graded readers used in the reviews could possibly have prevented the learners from using synonyms that they had acquired from the textbook. There are three reasons for this. Firstly, each reader contained vocabulary that would have been different from that provided in the textbook. Secondly, many (but not all) of the graded readers were part of a series featuring the same characters. Thirdly, the model text for the book review no doubt also had an effect on the choice of words, as many learners used certain formulaic expressions borrowed from the model text.

The learners had not learned many synonyms. Moreover, it is likely that they had not as yet consolidated their knowledge of the synonyms that they did know, or had not yet gained the confidence to make use of them. Nevertheless, the use of synonyms in the TBRs highlighted the need for more focus on this device in the lessons.

5.2.2.3 Conjunction

The findings on the use of conjunction, where the TBRs used fewer conjunctions but a wider variety, provides support for Hinkel (2001;125) in showing that NNSs tend to use conjunction more than NSs, and tend to rely more on sentence transitions such as *then* or *first* to make their texts cohesive. It was noted in §4.6.3.2 that, in some cases, the use of *and* and *because* to begin sentences when other options would have been more appropriate could be traced to certain sentences in the textbook, especially the use of dialog in the reading sections.

It was noted in Chapter 4 (§4.5.9.5) that the use of conjunction in the STs may also be explained in terms of learners' L1, particularly the use of *to* when *because* would have been

appropriate. These problems need to be kept in mind when accounting for errors in cohesion.

5.2.2.4 Reference

On one hand, the significantly greater use of reference in the STs is hardly surprising, since reference, particularly pronominal, was one of the first devices learned and was therefore very much part of the learners' active vocabulary. It is clear that the learners had, on the whole, mastered use of the pronominal even if individual errors did occur.

On the other hand, the significantly higher frequency of demonstratives supports the findings of Hinkel (2001), who noted that NNSs tended to use more demonstrative pronouns (§2.5.2.1) than NSs. This warrants further attention. The STs referred almost exclusively to the story or the book that was being reviewed (*this book* or *this story*, a device that was in fact modeled by the researcher), rather than to any items in the story. This suggests that they were, on the whole, not yet capable of referring to a wide variety of items using demonstratives. Although the learners may have been able to identify demonstratives in a text, they still needed more practice in using them in free writing. Nevertheless, the fact that some learners were able to use demonstratives correctly shows that, with the correct guidance (ideas about which are discussed in §5.5 below) the appropriate and effective use of these devices could be mastered at this level of learning. The learners clearly needed more practice not only in identifying demonstratives in texts but also using them in free writing exercises.

5.2.2.5 Repetition and superordinates

It was found that one reason for the lack of significant differences between the two corpora in collocation, repetition and superordinates lay in the fact that the TBRs featured a far greater variety of types than the STs. For example, in repetition, the TBRs repeated significantly more abstract nouns, stock phrases and adjectives than the STs, while the STs

repeated a significantly greater number of common nouns. Furthermore, with respect to superordinates, in the STs, 18 of the 44 superordinates identified referred to only one of two types, namely, *story* and *book*. By contrast, the most frequently used superordinate in TBRs (*man*) was used only four times, while all but two of the other superordinates were used only once.

These findings are consistent with previous studies. They support the findings of Abadiano (1995), who noted that most essays used lexical devices, particularly repetition, for cohesion. The use of repetition was also found to be the most commonly used device in Castro (2004), Mojica (2006) and Chen (2007).

However, the use of more different types (as opposed to the total number of tokens) in the TBRs is particularly interesting given that the textbook was working with a very limited vocabulary: it was limited to new words that were being introduced, plus words that had appeared previously. The new words that appear in the textbook are then listed at the bottom of the page. Had these texts been aimed at a more advanced reader, other types of reiteration such as synonyms or general words might have been more frequently employed. Throughout the book, the textbook writers were trying to reuse known words in order to reinforce them and to help learners to acquire them. Yet despite these limitations, the TBRs were able to use language naturally. How this was achieved was noted in the qualitative analysis, where certain cohesive devices such as direct repetition and conjunction were used far more sparingly and with more types than in the STs.

The limited variety of types in the STs suggests a general lack of awareness on the part the learners as to how cohesive devices can strengthen and enhance a text. Although this group of learners had been taught how devices such as superordinates were used in a text

to refer to something that had come before, they had not been given enough chance to practice these, nor had they been encouraged to create their own in compositions.

Working on a grammar problem on the one hand, and constructing original sentences to express one's ideas and opinions on the other, are two very different skills. Learners had only learned complex sentences four weeks before writing the essays. This may seem like a fairly substantial amount of time, but it needs to be borne in mind that the learners are only exposed to English for about seven hours per week, plus whatever time they spend studying at home. Relative clauses had also only been taught two or three weeks before. Although learners had had a great deal of practice in translating complex sentences, and other grammar exercises, they had not had nearly as much practice in using them freely. While the ability to use a wider vocabulary and to write extended sentences is not directly related to cohesion, this skill would possibly have allowed the writer to create more cohesive links between sentences with a wider variety of devices. This is because a wider vocabulary would firstly allow for the use of synonyms and superordinates, and conjunctions. Secondly, a wider vocabulary would mean that direct repetition could occur not simply among common nouns, as was largely the case with the STs, but with other parts of speech, particularly verbs and abstract nouns.

These findings revealed the different pace at which learners were progressing. Taking factors such as age, sex, intelligence and English background into consideration, the group seemed relatively homogenous, which appeared to lend support to Castro's (2004) findings that frequency of cohesive devices was not a factor in determining text quality among homogenous groups.

It was, however clear that the variety of cohesive devices was indeed a factor. There were clearly differences in the speed at which learners were absorbing and applying new

vocabulary and grammar structures, as well as writing skills. While some learners were able to write extended sentences, others could still only manage simple sentences. While some tried using synonyms or superordinates, others could merely repeat the same word a number of times. This was particularly evident in the qualitative analysis, where the writers of the better texts were able to use a variety of devices in creating cohesive chains, while the weaker texts tended to rely more on reference and repetition for cohesion.

There were a number of possible reasons for the different pace of learning. The first was motivation. While some learners clearly enjoyed English, and made an extraordinary amount of effort, others prioritized subjects such as maths and science over English studies, which they did not believe was entirely relevant to their future. There were also some learners who had not developed effective study habits in English. This is important, particularly in an EFL setting where learners are not immersed in the target language, and where consistency in study is crucial. Without regular study using a variety of strategies, these learners may also have had a smaller active vocabulary on which to draw. This means that they tended to rely on referencing or simple repetition, particularly in the central topic of the text.

5.2.3 Findings for Research Question 3

This section discusses the findings pertaining to Research Question 3: What patterns of reiteration occur in the TBRs?

Six features of good writing were identified in the TBRs. These are enumerated in §4.4 as follows:

1. A variety of cohesive devices used in cohesive chains
2. The opening sentence and repetition links
3. The central topic or theme
4. Clusters of repetition
5. Sentences of varied length
6. The final sentence

5.2.4 Findings for Research Question 4

This section relates to the findings of Research Question 4: How do the patterns of reiteration in the STs compare to those in the TBRs?

The similarities and differences between the selected STs and TBRs varied according to the quality of the STs. The more highly rated texts contained almost all of the features of good writing found in the TBRs, the average texts contained some but not all of the features, and the weaker texts contained few, if any. This offers some insight not simply into what cohesive devices were used, but also how their use contributed to the quality of a text.

The fact that some STs contained all six features found in the TBRs is important because it indicates not simply that some students at this level were capable of writing well-structured, interesting, informative texts, but because it also highlights the main features, as noted in the analysis of the TBRs, that make them so. If these features can be found in some learners, then it is possible that, with more practice, weaker learners may also acquire them. Indeed, the fact that the average texts contained some but not all of the features found in the TBRs suggests not only that there is a range of skills and abilities, but also that, with more practice, the weaker writers may well improve.

One of the reasons that the weaker texts contained few or none of the features found in the TBRs was, firstly, that many of the sentences were simply too short. In longer sentences

links and bonds may be created, while simple sentences can only convey limited information. . Secondly, the weaker texts used a more limited variety of cohesive devices. Thus, because these writers needed to keep repeating known information, they also had to rely more on reference and repetition.

The qualitative study supports the work of Okuda (2012) in highlighting the importance of not only introducing subtopics but also of linking these to the main topic, as well as elaborating on the topics introduced. The study also emphasizes the importance of clusters of cohesion, where ideas and topics were elaborated on, which lends support to Saldinha (1997), who noted how topic shifts were signaled through the breaks in lexical clusters. In the STs in this study, longer sentences provided more information which in turn created more links to other sentences, which not only reinforced the central theme of the text but also created links to the secondary topics.

5.2.5 Findings for Research Question 5

This section discusses the findings for Research Question 5: Which cohesive devices in the STs are used inappropriately or inaccurately? The study identified four main errors: exophoric referencing, redundant repetition, ambiguous referencing, and errors in conjunction.

With at least one error of cohesion in each of more than 50% of the texts, the need for more focus on cohesion in the classroom is clear. It was also revealed that 67% of the errors were committed by only 22% of the learners.

5.2.6 Findings for Research Question 6

This section discusses the findings pertaining to Research Question 6: What are the reasons for these errors?

L1 transference from Japanese may account for redundancy in repetition. For example, in Japanese, it is often preferable to repeat a person's name without using pronouns, in much the same way that Afrikaans speakers may avoid using pronouns when talking to someone senior to them. While pronouns exist in Japanese, their use is proscribed. For example, the second person pronoun *anata* is generally not used to address an unfamiliar person, particularly if that person is older or of a higher social rank. Furthermore, the third person pronouns *kare* (*he*) and *kanojo* (*she*) are not generally used when talking to or about someone with whom one is not very friendly or familiar, or in polite company. In these cases the person's name or title is used. Indeed, trying to decide how to address someone with whom one is unfamiliar is sometimes problematic even for native Japanese speakers. Whether or not this is the actual reason for this repetition is not entirely clear, but is a factor to consider. Similar findings were noted in Sadighi and Heydari (2012), where the cohesive errors found in Iranian texts could be traced to the grammatical rules, concepts and conventions found in the Persian language.

This is also highlighted in the work of McGee (2009) in his discussion of English written by Arabic-speaking learners of English, and Al-Khafafi (2005) in his study on the frequency of simple and complex repetition in Arabic and English texts. The present study noted in particular the overuse of the protagonists' names (as in the case with ST 5 (*The Scarf*) or ST 20 (*Greg's Microscope*)).

5.3 CONTRIBUTIONS OF THIS STUDY

The main aim of this study was to establish more precisely how the use of cohesive devices in learners' writing compared to their use in the learners' textbook. The important difference of this study from other studies is that it involved school learners at an elementary level of EFL. The first and most important finding of this study was that the participants were indeed

capable of constructing coherent essays, especially if provided with the tools to do so. Such tools included an overall structure with which to work, including a guide to what the text should contain.

One of this study's main contributions has been to highlight what such learners were capable of doing and what they were not capable of. It has proven to be a useful diagnostic tool in isolating a number of areas of weakness, such as the errors in cohesion, as well as areas of strength, namely, the fact that many learners in the study were capable of incorporating many of the features of reiteration identified in the textbook. Of course, the extent to which these findings may be applied to other populations is not entirely clear. However, the study has revealed some of the problems that elementary-level learners may encounter in writing.

By examining the cohesive devices used in the TBRs, the study revealed a number of options that are open to the elementary learner. For example, in addition to using reference to create cohesion, the learner at this level may also explore the use of superordinates and simple synonyms, not to mention different forms of words (e.g. *honesty* – *honest*) depending on the context of use.

In this study, the patterns of repetition clearly showed how the TBRs were anchored in a main topic, and how this anchoring then allowed the writers to move more easily to subtopics which themselves were anchored by one key repeated word. The fact that the stronger learners (for example, the writers of *Tom Sawyer* and *The Jumble Sale*) were capable of mirroring this pattern shows that learners at this level are capable of writing cohesive texts but need more guidance on how to do this.

This study has broken new ground in its focus on elementary-level junior high school learners in an EFL environment. One should not forget that, as learners of English as a

foreign language, the learners in this study do not have the privilege of being immersed in English in their everyday lives. Although their exposure to English is increasing, their main (and in some cases only) source of input remains their language teachers and the teaching materials provided. For this reason, the results of this study are of interest, not just to teachers in Japan, but to teachers in any EFL environment who would like to teach composition, but may feel that their learners are not yet ready to write extended texts.

On this note, in South Africa it is usually assumed that, because English is one of the official languages, most NNS learners are ESL learners. However, those in impoverished areas, where English is not a mother tongue or lingua franca, have little access to English, making them de facto EFL learners. For teachers and learners in such environments, this study will, I believe, offer some insights into how writing at the upper elementary and intermediate levels may be taught and improved.

I believe that this study will serve as a catalyst and a base for further research into writing by learners at this particular proficiency level, as well as offering a blueprint for teachers and learners to use when examining more closely the factors that affect writing at an elementary level of language learning. In addition, I firmly believe that this study will go some way to helping teachers and learners alike to become more aware of cohesion, to improve their compositions and, ultimately, find the satisfaction and fulfillment that can only be found in expression through writing.

5.4 LIMITATIONS

As has already been noted in this chapter, the data collected in this study were derived from a book review written under test conditions (but prepared beforehand). Because learners were given a model from which to work, there is the likelihood that this had an influence on the kind of language produced and the kinds of devices that were used.

Learners were still new to complex and compound sentences. Had the study been carried out a few months later when they were more skilled in building longer sentences, a different distribution of cohesive devices may well have been the result.

The essays written were relatively short. This too may well have had an impact on the findings, since longer essays may have forced longer sentences, which in turn may have increased the frequency of lexical devices in particular.

Finally, the learners were given time at home to prepare the essay which made it almost impossible to control for factors such as help from outside sources or the use of dictionaries to aid their work. It is possible, therefore, that the texts did not provide a completely accurate reflection of learners' writing ability. On the other hand, giving students time to prepare may have countered test anxiety as a factor in the quality of the texts.

5.5 RECOMMENDATIONS AND SUGGESTIONS FOR PRACTICAL APPLICATION

The findings in this study suggest a number of ways in which cohesion may be taught in the classroom.

An overabundance of simple sentences among elementary-level learners was the root cause of many of the cohesive errors. In order to help with sentence building, learners could be given exercises in which they expand on simple sentences, using the key WH&H questions: Who? What? When? Where? Why? How? Having been given a very simple sentence, such as 'The man went', learners are then asked to answer these questions in one long sentence. While such a sentence may not always be appropriate in essay writing, it does give students practice and confidence in creating longer sentences. Moreover, by answering the WH&H questions, the learners learn what is needed to create fuller sentences.

A similar activity involves a preliminary speaking exercise in which learners interview each other about a particular event, such as what they did at the weekend, or something that happened to them. They are then asked to collect as much information as possible using the WH&H questions. After this interview, learners write a few sentences about themselves or their classmates, including as much information as possible in one sentence. Learners can then analyze the work, noting the lexical and grammatical links between the sentences, as well as checking for problems with cohesion such as redundant repetition, exophoric referencing, and overuse of conjunction.

It was noted that ellipsis in this study had to be treated with some caution, since the use of dialog in many of the TBRs meant that there would naturally be more of these devices. However, this does not mean that ellipsis cannot or should not be discussed and taught at an elementary level. Moreover, it should also be noted that redundancy can often be remedied by the use of this ellipsis. One way of drawing attention to this device is by using what I would call 'reverse ellipsis'. This can be a very powerful tool for explaining certain grammar points. What this means is actually rewriting sentences and dialogs with the elided words included. This can help learners to understand the meanings of certain structures and clauses. For example, learners may struggle to decode the meaning of the following sentence:

I don't want to be taught by Mr. Smith, but by Mr. Jones.

(My own example)

This sentence is easily explained by writing two sentences out in full:

I don't want to be taught by Mr. Smith. I want to be taught by Mr. Jones.

(My own example)

One may then point out the words that have been repeated in the second sentence (*I want to be taught*), showing that they are in fact redundant. After crossing out these redundant

words and connecting the remaining fragments with *but*, the meaning of the sentence becomes clearer.

The use of superordinates in the STs was also an area where more work clearly needed to be done. While the number of superordinates used overall in the STs was not significantly different from the TBRs, most of these were simply 'book' or 'story'. A much greater variety of types was needed. One way in which to improve the use of superordinates may be by asking learners to find different ways to describe people or objects. For example, learners are given an article which repeats the name of a famous sports star. They are then asked to come up with different epithets, such as 'the 25 year-old', 'the current world champion', 'the center forward'. This would help to raise learners' awareness of the options that they have available to them. The use of synonym tables of words that have already been learned is useful in helping learners to summarize their work and to see the connection between the words that they have already learned. However, these should be retrospective tables that are expanded as synonyms appear, rather than long lists that are given in advance. Moreover, as McGee (2009:219) points out, one needs to warn against being simplistic in the use of synonyms and, wherever possible and appropriate, explain their different shades of meaning. Thirdly, one should provide practice in the use of hyponyms, which are essentially the opposite of superordinates (e.g. *the medical profession – doctor*). Fourthly, one needs to be aware of collocation, and how certain words go together with others, for example *ride* collocates with *bicycle*. Knowing which words commonly occur together can help to make language learning far less intimidating, since one becomes aware of the fact that in any given situation, the choice of words or expressions available is in fact limited.

The analysis of repetition patterns using the grid to track words and ideas highlighted the importance of the opening sentence in not only introducing the main theme, but also introducing words to other parts of the text. It showed how the main topic is reiterated, the

way in which clusters signal elaboration, and how ideas are linked, elaborated on, and reiterated using a variety of cohesive devices. Analysis of repetition patterns and lexical chains may also be used by learners themselves to discover how reiteration works. Learners could compare the patterns of reiteration in their textbook and in their own texts and note any differences. This could go a long way in raising awareness of cohesion

This study has shown that redundant repetition was another area needing to be addressed. This was supported by McGee (2009:219). One way of raising learners' awareness of redundancy is to write out a deliberately redundant text, and to ask learners to find ways of improving it, largely by deleting or substituting redundant repetition. Learners can also look at their own or each other's work, and note down which areas are repeated unnecessarily. In this case, work on the application of the known-new contract can go a long way to ensuring that information which is repeated occurs mainly in the first part of the sentence, with new information occurring in the verb and complement.

The findings clearly indicate that many learners seemed to understand what devices could be used, but did not have the vocabulary to use a variety of types accurately or correctly. To this end, we need to consider how vocabulary may be activated.

Awareness of the forms of words can help to improve cohesion. An ongoing chart which shows the different forms of words that have been studied is one possibility. However, as with synonyms, it must be stressed that such charts should not be for rote memorization, but rather to summarize words that have already been covered in the syllabus and to show how they change according to the part of speech. Students can also be given exercises in which they practice recognizing and using the different word forms in sentences and stories. As Chen's (2007) findings revealed, for many learners, the key to improving the use of lexical cohesion lay in ensuring that passive vocabulary became active vocabulary.

The work of Carell (1982) and Castro (2004) both emphasized the importance of activating learners' schemata. To this end, mind maps are useful in helping learners to group words according to different topics. They also help learners to summarize the words that they have already learned, and to see how any given topic usually limits one's options in terms of what words or expressions may be used. Thus the task of deciding which words to use in a given situation becomes less daunting. Mojica (2006) calls for lexical knowledge to be taken further. She believes that well-developed lexical knowledge can help to create lexically cohesive papers. She believes that students need to be encouraged to use more lexical items. These usually move from the specific to the general. Students need to be encouraged to use these in their writing wherever appropriate (McGee 2009:215).

In this study, L1 interference was also identified as one reason for cohesive errors, as also identified in the studies of Olateju (2006) Liu and Qi (2010), and Sadighi and Heydari (2012). One way that I have dealt with these problems has been to note the kinds of errors that occur in the learners' essays. I then compile a worksheet containing these errors (with content words changed in order to protect the identity of the student who committed them). In cases where an error was particularly prevalent among a number of students, I may include a number of examples (see Appendix 8 for one example of how to improve essays in this way). The learners then work through the worksheet and we look at how the errors could have been corrected.

Identification of cohesive errors is a useful diagnostic tool in establishing which errors were committed. A similar analysis could be carried out by a teacher with a group of students in order to establish which learners are struggling, or which errors are being committed by a larger number of learners. With this information it would then be possible to devise methods for addressing these problems. To this end, I have compiled worksheets containing sentences written by learners that are not necessarily erroneous but which could be

improved. This may often include problems with cohesion. We then work through these worksheets, either in groups or individually, and then look at how the work could have been changed to read more naturally.

Finally, following Nakao's (2009) study, more work needs to be done to develop a rubric for the assessment of student essays. Such a rubric should include cohesion as an important component in the assessment of student compositions.

5.6 SUGGESTIONS FOR FURTHER RESEARCH

Three suggestions for further study are offered.

Firstly, given that these learners had encountered very little English in their everyday lives, it is hard to tell whether the use of more authentic texts such as newspaper articles, magazine articles and advertisements would have enhanced their learning or would just have confused them. This could be investigated further.

Secondly, since the study did not focus on the relationship between high scores and the number of cohesive factors, it was difficult to discern with any level of accuracy whether such a relationship existed. The only assessment was that of the researcher. With more raters involved in the grading of all the essays, inter-rater reliability could have been established, thus strengthening the concept of a 'successful' text.

Thirdly, the qualitative analysis indicated that the patterns of lexis changed according to the quality of the text. Having focused on a small sample in order to focus on the qualitative aspects of the texts, each of the different features of a good text identified here could be investigated further, using a larger sample and further quantitative analysis, such as establishing the extent to which the features identified in the TBRs (§ 4.4) could be found in

a larger sample of student texts, and whether there was a statistically significant relationship between the quality of texts and the presence of these features.

5.7 CONCLUDING COMMENTS

This study has achieved its main goals. Firstly, it has shown precisely the type and frequency of cohesive devices that occur in both the TBRs. It has also established both quantitatively and qualitatively a number of reasons why the TBRs are, on the whole, more successful texts than the STs.

Through this study, it has become clear that the variety of cohesive devices used in writing is far more important to the quality of the text than the number of cohesive devices. The findings also suggest that many EFL learners at this level are capable of producing well-written, successful texts. However, what they needed in order to achieve this were more opportunities to consolidate their vocabulary by working with different cohesive devices, relating them to each other, and using them to construct their own original sentences. Clearly, learners in this study needed to develop a greater awareness of the lexis that they had at their disposal, of how, even with a limited vocabulary, words can be adapted, and how they can be reiterated and expanded upon without necessarily making the text boring or repetitive.

Given that some learners produced well-written, cohesive texts which made use of all the features of a successful text identified in the TBRs is proof that, with some time, effort and attention, even learners at an elementary level, are capable of improving their writing.

This study has not only identified some of the strengths and weaknesses of cohesion that occur in texts written by elementary-level learners in an EFL environment, but has also shown a number of ways in which the problems may be overcome. There is therefore no reason to doubt that the findings and suggestions in this study will help to make the

teaching and learning of writing more effective and more enjoyable, while helping elementary-level learners achieve their full writing potential.

BIBLIOGRAPHY

Al-Khafafi, R. 2005. Variation and recurrence in the lexical chains of Arabic and English texts. *Proznan Studies in Contemporary Linguistics* vol. 40:5–25. School of English, Adam Mickiewicz University, Poznan, Poland. [Online] Available at: <http://ifa.amu.edu.pl/psicl/files/40/01Al-Khafaji.pdf>

Abadiano, H.R. 1995 Cohesion strategies and genre in expository prose: An analysis of the writing of children of ethnolinguistic cultural groups. *Pragmatics* 5(3):299–324. International Pragmatics Association. [Online] Available at: <http://elanguage.net/journals/pragmatics/article/view/408/340>

Brown, J.D. 2004. *Understanding research in second language learning*. Cambridge UK: Cambridge University Press.

Castro, C.D. 2004. Cohesion and the social construction of meaning in the essays of Filipino college students' writing in L2 English. *Asia Pacific Review*, 5(2):215–255. [Online] Available at: <http://files.eric.ed.gov/fulltext/EJ720542.pdf>

Chen, Y. 2007. Lexical cohesion in Chinese college EFL writing. *Celea Journal*, 30(5):46–57.

Cook, G. 2004. *Discourse*. Oxford: Oxford University Press.

Coulthard, M. 1994. On analysing and evaluating written text, in Coulthard, M. (ed.), *Advances in written text analysis*. London: Routledge.

Cummins, J. 2008. BICS and CALP: Empirical and theoretical status of the distinction, in Street, B. and Hornberger, N.H. (eds), *Encyclopedia of language and education: Volume 2: Literacy*. 2nd edition. New York: Springer Science + Business Media LLC, 71–83. [Online] Available at: <http://daphne.palomar.edu/lchen/CumminsBICSCALPSpringer2007.pdf>

Cunningham, C. [sa]. Applying written discourse in a Japanese Class. ODL, Second Year at University of Birmingham. [Online] Available at: <http://www.cels.bham.ac.uk/resources/essays/Cunningham5.pdf>

Carrell, P. 1982. Cohesion is not coherence. *TESOL Quarterly*, 16(4):479–488.

Flynn, R.M. 2011a. *Progress in English 21 Book 2*. Osaka, Japan: Edec.

Flynn, R.M. 2011b. *Progress in English 21 Book 2 Teacher's Manual*. Osaka, Japan: Edec.

Fujieda Y. 2006. A brief historical sketch of second language writing studies. *Kyouai Gakuen Maebashi Kokusai Daigaku Ronshuu*, March. [Online] Available at: <http://www.kyoai.ac.jp/college/ronshuu/no-06/fujieda.pdf>

Fulcher, G. 1989. Cohesion and coherence in theory and reading research. *Journal of Research in Reading*, 12(2):146–163.

Gates, S. 2003. Inconsistencies in writing within the Japanese junior high school education system. *JALT Journal*, November, 25(2). [Online] Available at: <http://jalt-publications.org/archive/jj/2003b/perspectives.pdf>

Halliday, M.A.K. and Hasan R. 1976. *Cohesion in English*. London: Longman,

Hasan, R. 1984. Coherence and cohesive harmony, in J. Flood (ed.), *Understanding reading comprehension*. Newark, DE: International Reading Association, 181–219.

Hinkel, E. 2001. Matters of cohesion in L2 academic texts. *Applied Language Learning*, 12(2):111–132. Seattle University. [Online] Available at: <http://elihinkel.org/downloads/Cohesion.pdf>

Hoey, M. 1991. *Patterns of lexis in text*. Oxford: Oxford University Press.

Hofstee, E. 2006. *Constructing a good dissertation: A practical guide to finishing a master's, MBA or PhD on schedule*, Johannesburg, South Africa: EPE.

Hughes, A. 2003. *Testing for teachers*. Cambridge: Cambridge University Press.

Kimball, J. 1996. Writing on academic topics: Externalizing rhetorical processes in academic writing. *JALT Journal*. [Online] Available at: http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/15/0e/57.pdf

Kolln, M. 1999. Cohesion and coherence, in Cooper, C.R. and Odell, L. (eds), *Evaluating writing: The role of teachers' knowledge about text, learning and culture*. Urbana, IL: National Council of Teachers of English. [Online] Available at: <http://eric.ed.gov/?id=ED423540>

Li, H. 2011 Contrastive analysis of thematic progression patterns of English and Chinese consecutive interpretation texts. *Theory and Practice in Language Studies*, 1(6): 671-675.

Liu, D. 2000. Writing cohesion: Using content lexical ties in ESOL. *English Teaching Forum*, 38(1):28–35.

Liu, L. and Qi, X. 2010. A contrastive study of textual cohesion and coherence errors in Chinese EFL abstract writing in Engineering discourse. *Intercultural Communication Studies*, XIX:3. [Online] Available at: <http://www.uri.edu/iaics/content/2010v19n3/14LidaLiuXiukunQi.pdf>

Matsubara, Y. 2001. *Communicative writing: Teaching Japanese high school students to use writing to communicate in English*. Nakanishi Senior High School, Okinawa Prefecture. Opinion Papers. [Online] Available at: http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/17/03/a8.pdf

McGee, I. 2009. Traversing the lexical cohesion minefield. *ELT Journal*, July, 63(3). [Online] Available at: <http://eltj.oxfordjournals.org/content/63/3/212.abstract>

Mojica, L.A. 2006. Reiterations in ESL learners' academic papers: Do they contribute to lexical cohesiveness? *Asia Pacific Education Research*, December, 15(1):105–125. [Online] Available at: <http://www.dlsu.edu.ph/research/journals/taper/pdf/200612/mojica.pdf>

Monty Python. [Online] Available at: <http://www.montypython.net/scripts/petshop.php>

Muto, K. 2006. The use of lexical cohesion in reading and writing. *Journal of Foreign Language Studies*, Vol. 30:107–129. Nagoya University of Foreign Studies [Online] Available at: <http://ci.nii.ac.jp/naid/110003487239>

Nakao, T. 2009. Cohesion and coherence in writing samples of Japanese high school students: Implications for developing objective criteria in writing assessment. Paper delivered at Jacet Conference, 2009. [Online] Available at: http://www.n-junshin.ac.jp/jacet/2009/CONF_2009_ABS.pdf

Okuda, T. 2012. Patterns of lexical cohesion in writing samples of Japanese high school students: What can lexical cohesion tell us about the development of writing skills? *Association of Sophian English Teachers (ASTE) Newsletter*, Issue 70:17–32. Sophia University, Tokyo. [Online] Available at: <http://www.bun-eido.co.jp/aste/aste60.pdf>

Olateju, M.A. 2006. Cohesion in ESL classroom written texts. *Nordic Journal of African Studies*, 15(3):314–331. [Online] Available at: <http://www.njas.helsinki.fi/pdf-files/vol15num3/olateju.pdf>

Quinn, J. 2003. Cohesion: Form-focused and implicit feedback. *NVCB JLCC*, 5(1):41–48. [Online] Available at: <http://ci.nii.ac.jp/naid/110001046039>

Ramasawmy, N. 2004. Conjunctive cohesion and relational coherence in students' compositions. MA Applied Linguistics. University of South Africa, Pretoria, South Africa. [Online] Available at: http://books.google.co.jp/books/about/Conjunctive_Cohesion_and_Relational_Cohe.html?id=JE8fOAAACAAJ&redir_esc=y

Rayson, P. Log Likelihood calculator.

Available online at <http://ucrel.lancs.ac.uk/llwizard.html> [Accessed 15 May 2014]

Richards, J.C and Schmidt, R. 2002. *Longman Dictionary of Teaching and Applied Linguistics*. London. Pearson Education Limited.

Sadighi, F. and Heydari, P. 2012. Cohesion analysis of L2 writing: The case of Iranian undergraduate EFL learners. *Mediterranean Journal of Social Sciences*, May, 3(2). [Online] Available at:

http://www.mcser.org/index.php?option=com_content&view=article&id=306%3Acohesion-analysis-of-l2-writing-the-case-of-iranian-undergraduate-efl-learners&Itemid=63

Sardinha, T.B. [sa]. Lexical segments in text. Catholic University of Sao Paulo paper. [Online] Available at:

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.199.1034&rep=rep1&type=pdf>

Sasaki, M. 2000. Toward an empirical model of EFL writing processes: An exploratory study. *Journal of Second Language Writing*, 9(3):259–291. [Online] Available at: <http://sla.sjtu.edu.cn/thesis/Toward%20an%20Empirical%20Model%20of%20EFL%20Writing%20Processes--An%20Exploratory%20Study.pdf>

Seliger, H.W. and Shohamy, E. 2003. *Second language research methods*. Oxford: Oxford University Press.

Takagi, A. 2001. The need for change in English writing instruction. *The Language Teacher Online*, July, Issue 25.07 [Online] Available at: http://jalt-publications.org/tlt/issues/2001-07_25.7

Winter, E. 1994. Clause relations as information structure: Two basic text structures in English, in Coultard, M. (ed.), *Advances in written text analysis*. London: Routledge.

Witte, P. and Faigley, L. 1981. Coherence, cohesion and writing quality. *College Composition and Communication*, 32(2):189–204. [Online] Available at: <http://links.jstor.org/sici?sici=0010-96%28198105%2932%3A2%3AC189%3ACCAWQ%3E2.0.C%3B2-E>

Appendices

Appendix 1: Summary of cohesion and coding scheme in Halliday and Hasan (1976:333–338)

REFERENCE			R
1	Pronomials		1
(1)	singular, masculine	<i>he, him, his</i>	11
(2)	singular, feminine	<i>she, her, hers</i>	12
(3)	singular, neuter	<i>it, its</i>	13
(4)	plural	<i>they, them, their, theirs</i>	14
1(1-4)	functioning as:		
(a)	non-possessive, as Head	<i>he/him, she/her, it, they/them</i>	6
(b)	possessive, as Head	<i>his, hers, (its), theirs</i>	7
(c)	possessive, as Deictic	<i>his, her, its, their</i>	8
2	Demonstrative and definite article		2
(1)	demonstrative, near	<i>this/these, here</i>	21
(2)	demonstrative, far	<i>that/those, there, then</i>	22
(3)	definite article	<i>the</i>	23
2(1-3)	functioning as:		
(a)	nominal, Deictic or Head	<i>this/these, that/those, the</i>	6
(b)	place adverbial	<i>here, there</i>	7
(c)	time adverbial	<i>then</i>	8
3	Comparatives (not complete lists)		3
(1)	identity	<i>eg: same, identical</i>	31
(2)	similarity	<i>eg: similar(ly), such</i>	32
(3)	difference (<i>ie:</i> non-identity and dissimilarity)	<i>eg: different, other, else, additional</i>	33
SUBSTITUTION			S
1	Nominal substitutes		1
(1)	for noun Head	<i>one/ones</i>	11
(2)	for nominal Complement	<i>the same</i>	12
(3)	for Attribute	<i>so</i>	13

2	Verbal substitutes		2
(1)	for verb	<i>do, be, have</i>	21
(2)	for process	<i>do the same / likewise</i>	22
(3)	for proposition	<i>do so, be so</i>	23
(4)	verbal reference		24
3	Clausal substitutes		3
(1)	positive	<i>so</i>	31
(2)	negative	<i>not</i>	32
3(1-2)	substitute clause functioning as:		
(a)	reported		6
(b)	conditional		7
c)	modalized		8
(d)	other		9
ELLIPSIS			E
1	Nominal ellipsis		1
(1)	Deictic as Head		11
i	specific Deitic		1
ii	non-specific Deitic		2
iii	Post-deitic		3
(2)	Numerative as Head		12
i	ordinal		1
ii	cardinal		2
iii	indefinite		3
(3)	Epithet as Head		13
i	superlative		1
ii	comparative		2
iii	others		3
2	Verbal ellipsis		2
(1)	lexical ellipsis ('from right')		21
	i. total (all items omitted except first operator)		1
	ii. Partial (lexical verb only omitted)		2
(2)	operator ellipsis ('from left')		22
	i. total (all items omitted except first verb)		1

	ii. Partial (lexical operator only omitted)		2
	Note: Where the presupposed verbal group is simple there is no distinction between total and partial ellipsis; such instances are treated as 'total'.		
	Where it is above a certain complexity there are other possibilities intermediate between the total and partial as defined here; such instances are treated as 'partial'		
3	Clausal ellipsis		3
(1)	propositional ellipsis		31
i	total (all Propositional element omitted)		1
ii	partial (some Complement or Adjunct present)		2
(2)	modal ellipsis		32
i	total (all Modal element omitted)		1
ii	partial (Subject present) [rare]		2
	Note: Lexical ellipsis implies propositional ellipsis, and operator ellipsis implies modal ellipsis, unless all clause elements other than the Predicator (verbal group are explicitly repudiated.		
(3)	general ellipsis of the clause (all elements but one omitted)		33
i	WH- (only WH- element present)		1
ii	yes/no (only item expressing polarity present)		2

iii	other (other single clause element present)		3
(4)	zero (entire clause omitted)		34
3(1-4)	elliptical clause functioning as:		
(a)	yes/no question or answer		6
(b)	WH- question or answer		7
c)	reported' element		8
(d)	otherwise		9
	Note: Not all combinations of (1-4) with (a-D) are possible.		
CONJUNCTION	(items quoted are examples, not complete lists		C
1	Additive		I
(1)	Simple: (E/I)		11
i	additive	<i>and, and also</i>	1
ii	negative	<i>nor, and... not</i>	2
iii	alternative	<i>or, or else</i>	3
(2)	complex, emphatic: (I)		12
i	additive	<i>furthermore, add to that</i>	1
ii	alternative	<i>alternatively</i>	2
(3)	complex, de-emphatic	<i>by the way, incidentally</i>	13
(4)	apposition: I		14
i	expository	<i>that is, in other words</i>	1
ii	exemplificatory	<i>eg, thus</i>	2
(5)	comparison: (I)		
i	similar	<i>likewise, in the same way</i>	1
ii	dissimilar	<i>on the other hand, by contrast</i>	2
2 Adversative			2
(1)			21
i	simple	<i>yet, though, only</i>	1
ii	+ 'and'	<i>but</i>	2
iii	emphatic	<i>however, even so, all the same</i>	3
(2)	contrastive (avowal): I	<i>in (point of) fact, actually</i>	22

(3)	contrastive: (E)		23
(i)	simple	<i>but, and</i>	1
(ii)	emphatic	<i>however, conversely, on the other hand</i>	2
(4)	correction: (I)		24
(i)	of meaning	<i>instead, on the contrary, rather</i>	1
(ii)	of wording	<i>at least, I mean, or rather</i>	2
(5)	dismissal: (I)		25
i	closed	<i>in any / either case</i>	1
ii	open-ended	<i>in any case / anyhow</i>	2
3. Causal			3
(1)	general: (E/I)		31
(i)	simple	<i>so, then, therefore</i>	1
(2)	emphatic	<i>consequently</i>	2
-2	specific (E/I)		32
i	reason	<i>on account of this</i>	1
ii	result	<i>in consequence</i>	2
iii	purpose	<i>with this in mind</i>	3
(3)	reversed causal (I)		33
(4)	causal, specific: (I)		34
i	reason	<i>it follows</i>	1
ii	result	<i>arising out of this</i>	2
iii	purpose	<i>to this end</i>	3
(5)	conditional (E/I)		35
i	simple	<i>then</i>	1
ii	emphatic	<i>in that case, in such an event</i>	2
iii	generalized	<i>under the circumstances</i>	3
iv	reversed polarity	<i>otherwise, under other circumstances</i>	4
(6)	respective: (I)		36
i	direct	<i>in this respect, here</i>	1
ii	reversed polarity	<i>otherwise, apart from this, in other respects</i>	2

4. Temporal			4
(1)	simple: (E)		41
i	sequential	<i>then, next</i>	1
ii	simultaneous	<i>just then</i>	2
iii	preceding	<i>before that, hitherto</i>	3
(2)	conclusive: (E)	<i>in the end</i>	42
(3)	correlative: (E)		43
i	sequential	<i>first ... then</i>	1
ii	conclusive	<i>at first / originally / formerly... finally / now</i>	2
(4)	complex: (E)		44
i	immediate	<i>at once</i>	1
ii	interrupted	<i>soon</i>	2
iii	repetitive	<i>next time</i>	3
iv	specific	<i>next day</i>	4
v	durative	<i>meanwhile</i>	5
vi	terminal	<i>until then</i>	6
vii	punctiliar	<i>at this moment</i>	7
(5)	internal temporal: (I)		45
i	sequential	<i>next, then</i>	1
ii	conclusive	<i>finally, in conclusion</i>	2
(6)	correlatives: (I)		46
i	sequential	<i>first... next</i>	1
ii	conclusive	<i>in the first place ... to conclude with</i>	2
(7)	here and now: (I)		47
i	past	<i>up to now</i>	1
ii	present	<i>at this point</i>	2
iii	future	<i>from now on</i>	3
(8)	summary: (I)		48
i	summarizing	<i>to sum up</i>	1
ii	resumptive	<i>to resume</i>	2
5. Other (‘continuative’)		<i>now, of course, well, anyway, surely, after all</i>	5
6. Intonation			6

(1)	tone		61
(2)	tonicity		62
LEXICAL			
1	same item		1
2	Synonym or near synonym (incl hyponym)		2
3	Superordinate		3
4	General' item		4
5	Collocation		5
1-5	having reference that is		
(a)	identical		6
(b)	inclusive		7
c)	exclusive		8
(d)	unrelated		9

Appendix 2: Writing assignment

Name: _____ Class _____ Number _____

BOOK REVIEW

Due: 期末 test

Answer some of these questions about the book. (You do NOT NEED to write full sentences).

When you have finished, write a short book review. The review should be more than 60 words.

1. What is the name of the book? _____
2. Where does the story happen? _____
3. Who are the main characters? _____
4. Describe (説明する) the main characters.

5. What happens in the story? (Do not give the ending)

6. Why do you like this book OR Why do you **not** like this book?

7. What was your favorite part of the story? Why did you like it?

8. Did you learn something from this story? If 'yes', what did you learn?

Here is an example of a book review

Book review: The Monkey's Paw

'The monkey's paw' **is a story about** a very close, but simple family that receives a magic monkey's paw. If a person makes a wish, the paw will give them the wish, but it will also bring bad luck. The family is not rich, but they are happy. What will happen if they use the monkey's paw to make a wish?

The story is a very sad, but I enjoyed reading it. There were many interesting surprises. I think that the family members are a bit foolish. But maybe we all think the same as these people. This story showed me that it is better to be happy with what you have than to always want more.

Here are some useful expressions (表現)

'The Wrong Trousers' is about a ___ who ___

I thought the book was ___, because ___

I found the book very (exciting), because ___

My favorite part of the book is when _____

I would recommend this book, because___

I would not recommend this book, because___

This book taught me / showed me that ___

Appendix 3: Letter of consent

A request from Mr. Coombe

I am busy carrying out a study of student's essays as part of my Masters studies. I would therefore like to ask permission for your randomly selected essay (Book review, written in the July Kimatsu (term-end exam) to be used as a source of data.

Focus of the study:

The study that is being carried out is a comparison of **cohesive devices** found in the student's essays with some of the reading passages found in the text book, Progress 21.

What is cohesion?

Cohesion is the way in which certain words refer back to other words in a text, and help keep the text together. For example:

John was born in New York. **He** always loved **this exciting city**.

In the second sentence, 'He' refers back to 'John', while 'this exciting city' refers back to 'New York'. 'He' and 'this exciting city' are two types of cohesive device. The way in which people use cohesive devices can affect their writing style.

Promises by Mr. Coombe

1. Because this is a random sample, except for some minimal requirements, selection or non-selection into the study is in no way a reflection of the quality of your writing.
2. Your essay will be given a different number in order to protect your privacy. Your name, personal history, or student number will not at any stage be used.
3. If you wish, you will also be told about the results of the study, and will be allowed to see it when it is finished.

Letter of consent

By signing this document, I am giving permission for Deneys Coombe to use the above essay as part of his thesis.

Signed: _____ Class: _____ Student no: _____ Date: _____

Appendix 4: Sample of student text coding

Text 24: 86 words

Robin Hood

“Robin Hood” is an adventure story.

Anneena, Biff and Kipper played at Biff’s home.

Suddenly, the magic key which Kipper had took them to a wood and then they met Robin Hood there.

But the Sheriff grabbed Robin Hood and the children.

Will the children be able to come back to Biff’s house?

I would recommend this book because I was exciting when I read this book.

And you can read this book because it’s not too difficult to read.

Please try to read “Robin Hood”.

The sample above was coded as follows.

Note: The word in bold denotes an error.

Sentence No.	No. of ties	Cohesive Item	Type	Distance	Presupposed Item
2	1	Anneena, Biff, Kipper	L1		exophoric - characters in this series
3	3	the magic key	L1		exophoric - characters in this series
		Kipper	L1	0	Kipper
		them	R14.6	0	Anneena, Biff, Kipper
4	5	But	C21.2	0	s(3)
		the Sheriff	L5	0	Robin Hood story
		Robin Hood	L1	0	Robin Hood
		the	R23.6	0	them - the children - Anneena, B, K
		children	L3	0	them - the children - Anneena, B, K
5	4	the	R23.6	0	the children
		children	L1	0	children
		come back	L5	N1	took
		Biff’s hous	L1	N2	Biff’s House
6	2	this X 2	R21.8	N4	story
		exciting (excited)	L5	N4	story
		this	R21.8	0	book
		book	L1	0	book
		read	L5	0	book
8	2	read	L1	0	read
		"Robin Hood"	L1	0	this book - Robin Hood

Appendix 5: Reading texts from Progress in English Book 2 (Flynn 2010a)

Text A: - EASTER AT GRANDMA'S 152 Words

Tom and Mary Green usually stay with Grandma and Grandpa White during the spring holidays. It's Easter Sunday.

It's a beautiful spring day.

Last night Grandma painted a dozen Easter eggs for her grandchildren.

Early this morning she hid them in the backyard.

Tom and Mary are trying to find them.

"Here's one!" Tom shouts.

"Under this bush!

And here's one more!"

Mary looks and looks but can't find any eggs.

Tom finds two more near the gate.

He sounds very happy, but Mary is beginning to feel sad.

"Grandma," she says, "I can't find any!

Where did you hide them?"

"Maybe you should look in the doghouse.

Maybe the Easter Bunny left one there."

Grandma laughs.

"In the doghouse?!"

Mary looks in the doghouse and finds half a dozen pretty eggs right inside the entrance!

"Hey, look!" she shouts.

Now Mary looks happy, too.

Spring brings everyone new life and hope.

Text B: - A BEAUTIFUL SYMBOL OF FRIENDSHIP (p10)

Early in spring, cherry blossoms bloom along the Potomac River in Washington.

There are hundreds of trees along the river.

They look very beautiful, not only during the day but also at night.

Where did all these cherry trees come from?

Actually, they came from Japan!

They were a special gift from Japan to America.

At the beginning of the 20th century, the mayor of Tokyo sent the American people this wonderful gift from the Arakawa River.

He wanted to promote friendship between Japan and the United States.

Fifty years later, America returned the favor.

The old trees along the Arakawa River were dying, so America sent Tokyo many young cherry

trees from the Potomac River.

This gift from America made the Arakawa River beautiful again.

Not only in Tokyo but also in Washington, everyone looks forward to the cherry blossoms every year.

These trees are still a beautiful symbol of Japanese-American friendship.

Text C: - THE BIG APPLE

People call New York “the Big Apple,” but this is a rather new name for New York.

When the Dutch started a colony there in 1624, they called it “New Amsterdam.”

Two years later they “bought” Manhattan Island from Native Americans for almost nothing.

This Dutch colony became a very important city.

Then, in 1664, the English took it from the Dutch.

They named it “New York” because York was a famous old city of England.

At the beginning of the 20th century, someone called New York “the Big Apple.”

Why?

No one knows.

What a strange name for a big city!

But people still call it “the Big Apple” today.

The people of New York are from every part of the world.

They come not only from western and eastern Europe but also from Africa and Asia, from Central and South America.

The Statue of Liberty welcomes everyone.

While you are in America, be sure to visit New York and enjoy the taste of this Big Apple!

Text D: - THE “NEW WORLD” (p22)

In fourteen hundred and ninety-two / Columbus sailed the ocean blue.

Every American knows that year and that name.

Christopher Columbus was a brave Italian sailor.

He had a dream.

“The earth is round,” he said, “like a huge ball.

I’m going to sail west and reach India.”

“You’re crazy,” everyone said.

“The earth is flat.

You’ll fall off if you sail west.

You have to go east to reach India.”

But in 1492 Columbus was able to sail west because the Queen of Spain believed him and gave him three ships.

Columbus sailed across the Atlantic Ocean and reached some islands between North and South America.

“This is India!” he said, and he called the people “Indians.”

When people in Europe heard about this “new world,” many explorers decided to go there.

One of these explorers was Amerigo Vespucci.

He explored the long coast of South America and made maps of it.

People began to call the New World “America.”

In 1607, a group of people from England arrived on the east coast of North America.

They started a colony and called it “Jamestown” because James was the name of their King.

These colonists found life in the new world difficult.
When they arrived in spring, they were able to hunt for rabbits, turkeys and deer.
They planted some vegetables, too, but when winter came, they did not have enough food.
Many of the colonists got sick and died.
Some stayed in America, but many returned to England.
Others came and started new colonies along the coast.
Thirteen of these colonies became the first thirteen states of the U.S.
Count the number of stripes on the American flag.
There are thirteen.
They stand for these first thirteen states.
There are also fifty stars.
They stand for the number of states in America today.

Text E: - POCAHONTAS, A BRAVE AMERICAN GIRL (p30)

One of the colonists in Jamestown was Captain John Smith.
He was a brave leader.
When the colonists found life in the New World difficult, he asked, “Which do you like better, this new world or England?”
Some went back to England, but many stayed in the colony.
They said, “Life here is a lot better than our old life in England.”

One day some Native Americans captured Captain John Smith and took him to their chief.
When the chief saw Captain Smith, he said, “This man is our enemies’ chief.
Kill him!”

“No!” said Captain Smith, “We are not your enemies.
We want to become your friends.
This is a big land.
We can live together in peace.”

The chief repeated, “These people kill our deer.
They steal our land.
Kill him here and now!”

Just then, the chief’s daughter, Pocahontas, ran to Captain Smith.
She put her arms around him and cried out, “No, Father, please!
Please don’t kill him!”

Which do we need more, friends or enemies?

If you save this man now," she said, "he and his people will surely become our friends."

Smith promised to send the chief some guns and the chief sent him back to the colony. Pocahontas not only saved Captain Smith's life but also promoted friendship between her people and the English.

Later, Pocahontas married an English colonist, John Rolfe.

She went to England with him and their son, Thomas.

The King of England welcomed this brave American girl and she became very popular among the English.

Unfortunately, she became ill and died while she was in England, but she lives on in history.

Text F: - THE BOSTON TEA PARTY (p34)

For more than a century the colonists were dependent on Britain for many goods.

They had to pay taxes on these British goods and then they had to pay more taxes in order to unload the ships.

These taxes were making the British richer and the colonists poorer.

The colonists were beginning to ask, "Which is better, dependence on Britain or independence?"

One night in December, 1773, the colonists finally decided to do something about this.

Mrs. Hewes: George, where were you?
I was worried about you.

George Hewes: I went to a tea party.

Mr. Hewes: You went to what?!
In the middle of the night?

George Hewes: Yes.
We made a huge cup of tea, as huge as Boston Harbor.

Mrs. Hewes: Was it better than my tea?

George Hewes: Well, maybe it wasn't as good as yours, but it was sure bigger.

Mrs. Hewes: George, what are you saying?

George Hewes: You know those three big ships in the harbor?
They're full of tea.

But we can't unload them if we don't pay a special tax.

Mrs. Hewes: Taxes!
Taxes on this, taxes on that!
If we're not careful, we'll have to pay taxes in order to eat breakfast
or change our clothes.
The King's men don't care about us.
They're just using us to make money for England.

George Hewes: Right!
So we painted our faces and dressed like Mohawks.
We got onto the ships, opened 342 big boxes of tea and threw all the tea
into
the harbor.

Mrs. Hewes: George, you didn't!

George Hewes: We sure did!
So if you want a cup of good British tea, go to the harbor.
It'll taste a little cold and salty, but there's enough for everyone!

Mrs. Hewes: George, this will mean war!

George Hewes: Right again, dear!
It will be a war for independence.

Text G: - PUT YOUR "JOHN HANCOCK" HERE

Tom took Jiro to the front office of the school to sign some papers.

The woman in the office said, "Just put your John Hancock here."

Jiro did not understand and looked at Tom.

"That means 'Sign your name,'" Tom explained.

"Do you have a two-dollar bill?" Tom asked the woman in the office.

"I want to show him the picture on the back."

"Two-dollar bills are not very popular, you know," the woman said and went to look.

"They made them in 1976 to celebrate America's two hundredth birthday, but for some reason people don't like to use them.

OK.

Here's a two-dollar bill."

She brought the two-dollar bill and showed Jiro the back side.

“This is ‘The Signing of the Declaration of Independence,’” she explained.

“It’s one of the most famous American paintings.

At the beginning of July, 1776, delegates from all the colonies were meeting in Congress in Philadelphia to talk about independence.

Thomas Jefferson wrote the Declaration and Congress discussed it.

Some of the delegates were beginning to feel more and more nervous.

The Declaration meant independence from Britain and probably war!

“Finally, on July 4, the delegates accepted the Declaration.

The President of Congress signed his name very carefully and in big letters—John Hancock. Four days later, the Liberty Bell rang to tell the world about America’s independence from Britain.”

“And so,” Tom added, “when people want your signature, they sometimes say ‘Put your John Hancock here.’”

Text H: - AMERICAN FOOD (p46) 262 words

Americans eat many kinds of food from many different countries, but what kind of food is typically "American"?

Most Americans will probably say sandwiches, hamburgers or hot dogs, with French fries and salad, for lunch.

For dinner in the evening many will choose soup, salad, meat, potatoes, vegetables and dessert.

One Saturday Tom takes Jiro to a food court for lunch.

Jiro: Wow! What a big place! You can find almost anything here, can't you?

Tom: Yeah. But you're probably going to order a hamburger, aren't you?

Jiro: Just a minute.

Yeah, I'll have a hamburger, two pieces of pizza, a tuna sandwich and strawberry ice cream.

Tom: Hey! Take it easy!

You're not going to have all that for lunch, are you?

Jiro: Yeah.

Maybe you're right.

I'll just have two cheeseburgers, two pieces of pizza and a banana split.

Tom: You'll get fat if you eat like that.

In fact, you'll get fatter and fatter and won't be able to get into the plane when you have to go back to Japan.

Jiro: OK, OK!

Just one cheeseburger with French fries, a hot dog with ketchup and mustard and a large cola float.

Of course, everything was much larger than in Japan.

The cheeseburger was almost bigger than the plate, the hot dog was a foot long, and the cola float was a giant.

Jiro was able to finish everything, but he didn't want any dinner that evening.

The next time he didn't order so much.

He didn't want to become the fattest boy in New York.

Text I: - THE WHITE HOUSE (p54) 281 words

Tom and Jiro are going to visit the White House, which is in the middle of Washington .

They are looking at pamphlets which have all kinds of information about the White House.

Tom: According to this schedule, the President is going to be home this afternoon.

Jiro: Hey, let's go and visit him.

Tom: Don't be silly.

I can't just phone and say, "There's boy here that came all the way from Japan to see you."

Jiro: Why not?

It'll probably make him very happy.

Tom: Read your pamphlet, Jiro.

(They read for a while.)

Jiro: "The White House has 132 rooms."

That's bigger than the Whites' house in Kobe!

We have just four rooms in our house—two bedrooms, a living room and a kitchen.

There's a bathroom, of course, and a toilet, too, but only one!

Tom: The White House has a Green Room, a Blue Room, a Red Room, an East Room and a formal dining room with chairs for 140 guests.

Here's a picture of the Green Room.

Jiro: Everything's green!

The walls, the carpet, the chairs!

How yucky!

Tom: Look at this.

“George Washington was the only President that never lived in the White House.”

Jiro: Maybe he didn't like it.

All those strange rooms—green, blue, red.

Tom: There was no White House when he was President.

In fact, George Washington became the first President of the U.S. in New York!

Jiro: Oh, really? I didn't know that.

Tom: Look at this.

(He shows Jiro a 20-dollar bill.)

Twenty-dollar bills have a picture of the White House on the back.

Jiro: Hey, this is nice!

Can I have it?

Tom: No way!

Text J: - ABRAHAM LINCOLN

Abraham Lincoln was President of the U.S. during the Civil War between the northern and the southern states.

Lincoln was born on February 12, 1809, in a little log house and grew up on a farm.

The Lincoln family was very poor.

They were not able to send "Abe" to school or buy him books, but he borrowed books and read them again and again.

Abe studied hard and became a successful lawyer.

He was famous for his honesty.

They called him "Honest Abe."

One day, while he was still young, Abe saw something that he never forgot.

He was in New Orleans on business.

Some rich merchants were buying and selling black people from Africa.

Slavery was very common in the southern states in those days.

The slaves whom the white people bought worked on the farms in the South.

"I'll never see my children again!" one African woman was shouting.

"This man sold them to a rich white man who lives far from here."

The merchant who was selling the slaves saw Abe and asked, "You, sir, do you want to buy this woman?"

She works very hard."

"Of course not!" answered Abe.

"You can't sell people like dogs or potatoes!"

"Come on, come on!"

They're not people.

They're just slaves," said the merchant.

Lincoln was getting angrier and angrier.

"Someday, if I can, I'm going to stop this terrible business," Abe said.

The merchant laughed and continued his business.

Many years later, when Lincoln was President, he did not forget his promise.

In 1863, during the terrible Civil War, he kept his promise and freed all the slaves.

Text K: - MARTIN LUTHER KING (p66)

"I have a dream."

The man who said these famous words was a Protestant minister.

His name was Martin Luther King.

He spoke on the steps of the Lincoln Memorial, at the feet of the great statue of Abraham Lincoln.

The year was 1963.

It was a hundred years after Lincoln freed the slaves, but the grandchildren of the slaves were still

not really free.

The South was not a nice place for black Americans.

They did not have all the rights that white people had.

Black children and white children went to different schools.

Buses had seats in the front for white people and seats in the back for blacks.

Martin Luther King wanted to change this, but he did not want to use violence.

Of course, as a Christian minister he knew the life and words of Jesus Christ.

He also knew the life and words of Mahatma Gandhi.

According to both of these great men, love is more powerful than hate.

King respected these men and their beliefs.

He became a powerful speaker and a brave leader.

He led many peaceful protests against unfair laws.

One day, in December, 1955, Rosa Parks, a black woman, got on a bus in Montgomery, Alabama.

She was very tired at the end of a long day of work.

There were no empty seats in the back of the bus, so she took a seat in the front of the bus.

The bus stopped, but Rosa Parks did not move.

The police came and arrested her and took her to jail.

Black leaders in Montgomery gathered at Martin Luther King's house to talk about this.

"No black people will ride on buses," they decided.

Black people walked to work for a year.

The bus companies lost a lot of money.

Many people around the world noticed this silent protest.

"I have a dream," King said, "that one day, in Alabama, little black boys and black girls will be able to join hands with little white boys and white girls and walk together as sisters and brothers.

I have a dream today!"

Martin Luther King continued to work for the rights of African Americans.

He received the Nobel Peace Prize in 1964.

Four years later someone shot and killed him.

Martin Luther King died, but his dream lives on.

TOTAL 2849/6482 words

Appendix 6: Raw scores and complete tabulations

Table A6.1: Raw Scores: Total number of cohesive devices in student texts (grammatical)

Note: a space with no score can be read as 0.

	Grammatical Devices				
Code:	C	E	R	S	
Text No:	Conjunction	Ellipsis	Reference	Substitution	Total
1	2	1	8		11
2	0		4		4
3	1	4	7		12
4	2		12		14
5	0		4		4
6	1		6		7
7	1		8		9
8	2		6		8
9	0		3		3
10	5		10		15
11	0		8		8
12	5	2	11		18
13	0		2		2
14	3		5		8
15	0		13		13
16	0		6		6
17	0	1	6		7
18	2		8		10
19	0		3		3
20	0		7		7
21	3		8		11
22	1	1	8		10
23	1		7		8
24	3		5		8
25	1		4		5
26	1		7		8
27	6		22		28
28	1		6	1	8
29	1		9		10
30	0		3		3
31	5	2	11		18
32	1		9		10
33	3		6		9
34	2		11		13
35	3	1	8		12
36	0	1	5	1	7
Total	56	13	266	2	337
Mean	1.556	0.361	7.389	0.056	9

Table A6.2: Total number of cohesive devices in student texts (lexical)

Lexical Devices					
L1	L2	L3	L4	L5	
Repetition	Synonyms	superordinates	General Item	Collocation	Total Lexical
4		1		2	7
0		2		1	3
9		1		6	16
12	1	1		4	18
11				2	13
4		2		3	9
6	1			6	13
22	1			0	23
4		2	1	3	10
14		1		2	17
6		1		0	7
15				9	24
11				6	17
8		4		2	14
7				4	11
2		4		5	11
7		2		6	15
10		2		4	16
9		5		3	17
9				3	12
15				9	24
10		2		4	16
7	1	1		2	11
9		1		3	13
5		1		5	11
24				5	29
12	2	1		2	17
4		2		6	12
6		2		4	12
6	2			6	14
12		1		7	20
6		1		5	12
6		3		3	12
9	2	1		7	19
8		1		8	17
9				5	14
318	10	45	1	152	526

Table A6.3: Number of mistakes in each text, and total number of cohesive devices including mistakes

Text No:	Total excl mistakes	Mistakes	Total incl mistakes
1	18	1	19
2	7		7
3	28		28
4	32		32
5	17		17
6	16		16
7	22	4	26
8	31		31
9	13	3	16
10	32	4	36
11	15		15
12	42	1	43
13	19		19
14	22	3	25
15	24	2	26
16	17		17
17	22		22
18	26	1	27
19	20		20
20	19		19
21	35	3	38
22	26	2	28
23	19	4	23
24	21		21
25	16		16
26	37	3	40
27	45	2	47
28	20		20
29	22		22
30	17		17
31	38	1	39
32	22		22
33	21		21
34	32	1	33
35	29		29
36	21	1	22
	863	36	899
	23.972	1.000	24.972

Table A6.4: Total number of cohesive devices for textbook readings (grammatical)

Code:	C	E	R	S	
Text No:	Conjunction	Ellipsis	Reference	Substitution	Total
A	2	5	6	1	14
B	3		11		14
C	2	2	8		12
D	1	5	23		29
E	0	3	28		31
F	2	5	15	3	25
G	3	3	11		17
H	4	4	16	1	25
I	1	10	19		30
J	1		17	1	19
K	0	1	19		20
Total	19.00	38.00	173.00	6.00	236

Table A6.5: Total number of cohesive devices for textbook readings (lexical)

Code:	L1	L2	L3	L4	L5	
Text No:	Repetition	Synonyms	superordinates	General Item	Collocation	Total
A	22	2	3		8	35
B	33	4	2		9	48
C	11		5		6	22
D	36	6	7		24	73
E	37		6		11	54
F	23	4			8	35
G	22	2			8	32
H	16	1	2		21	40
I	19	3	2		15	39
J	23	6	6		12	47
K	32	5	9		21	67
Total	274.00	33.00	42.00	0.00	143.00	492

Appendix 7: Superordinates found in the TBRs

TBR	Superordinate	Referent
A	eggs	Easter eggs
B	America	Washington
B	gift	Cherry trees
B	Japan	Tokyo
B	trees	cherry trees
C	big city	New York
C	Dutch colony	Manhattan
C	everyone	people
C	Manhattan	New York
D	name	Christopher Columbus
D	dream	sail west
D	explorers	sailors
D	food	rabbits, turkeys
D	go	sail
D	New World	islands near America
D	number	fifty
E	brave ... girl	Pocahontas
E	enemies	colonists
E	man	Captain John Smith
E	New World	Jamestown
E	people	people
H	countries	American
H	place	American food court
I	a boy	Jiro
J	days	When Lincoln was a young man
J	southern states	New Orleans
J	the South	New Orleans
J	business	buying and selling slaves
J	promise	stop slavery



K	beliefs	Love is more powerful than hate
K	black	this (unfair laws)
K	Christian minister	Martin Luther King
K	famous words	'I have a dream'
K	four years later	1968
K	great men	Christ, Gandhi
K	men	men (Christ, Gandhi)
K	silent protest	against the bus laws

Appendix 8: An example of exercise on improving writing

Name _____ Class _____ Number _____

Let's make our writing better ... but how?

Here is a simple essay. It is ok, but it is not very interesting, is it? Why do you think so?

I like summer. I play in summer. Summer is hot. I eat fruit in summer.

Summer is nice. Summer is fun. So I like summer.

Here are some useful Points:

1. *Don't say the same thing too often.*
2. *Don't make all sentences the same length (長さ) (See 1).*
3. *Use different kinds of grammar when you can.*
4. *Don't use words and sentences that you do not need.*
5. *Do not say things that everyone already knows.*
6. *Keep it simple, but not too simple..*
7. *Answer the 'questions: **Who? Where? What? Why? When? and How?***
8. *Say **new** things.*

Ex: I like spring, because the cherry blossoms **make me happy.**

9. *Ask yourself: **'Who will be reading this?'** What do they know? What don't they know?*

If they know it, I don't need to say it.

If they don't know it, I need to explain (説明する).

Try to make these sentences better:

1. I like summer, because summer is hot, but winter is cold. So I like summer.

2. This summer, I want to swim in the sea, because I like swimming very much. I also want to go to Okinawa, because I like Okinawa very much.

3. I visit my grandparents. I like my grandparents. I do many things when I stay at my grandparents' house.

Appendix 9: Letter of ethical clearance



MEMORANDUM

From: Prof FA Kalua
Chair: Higher Degrees Committee,
Department of English Studies

To: Mr Coombe
MA, student no: 749 733 4

Date: 26 June 2014

SUBJECT: COMPLIANCE WITH RESEARCH ETHICS OF THE UNIVERSITY OF SOUTH AFRICA

This letter confirms that Mr Coombe has complied with the requirements of the Higher Degrees Committee in the Department of English Studies. His proposal has been approved and he is now working on the research study on The Cohesion Factor: A Study of Japanese Junior High School Writing.

The department confirms the following:

- Mr Coombe is a registered MA student in the Department of English Studies at UNISA
- He is expected to work closely with his supervisor, Mrs R Scheepers.
- He is aware of the Unisa Ethical Research policy and is expected to adhere to the policy requirements.



Prof FA Kalua

Date: 08/07/14

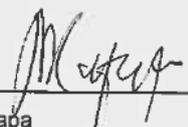
Chair: Higher Degrees Committee



Mrs R Scheepers

Date: 7.07.14

Supervisor: English Studies



Prof L Rafapa

Date: 08/07/14

Chair of Department: English Studies