

Exploring the Role of Model Essays in the IELTS Writing Test:  
A Feedback Tool

Makoto Abe

The University of Queensland

Supervisor: Dr. Noriko Iwashita

Submitted in partial fulfilment of the requirement of the Degree of Master  
of Arts in Applied Linguistics, The University of Queensland

May 2008

### **STATEMENT OF AUTHORSHIP**

I declare that this thesis does not contain material which has been accepted for the award of any other degree or diploma in any university, nor does it contain material previously published or written by any other person, except where due reference is made in the text of the thesis.

Makoto Abe

## ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor, Dr Noriko Iwashita, for her patient support, valuable suggestions, and expert guidance throughout all stages of this project.

I also would like to express my deep appreciation to Mr. Paul Forster, Ms. Melika Stannard, and Mr. Satoshi Nagaba for helpful advice and invaluable support during the data analysis.

My appreciation also extended to Mr. Noboru Sakai and Mr. Ken-ichi Hashimoto for their thoughtful support on proof-reading and insightful suggestions on this thesis.

I wish to express my gratitude to my friends at The University of Queensland. In particular, a sincere note of thanks goes to, Ms. Mayuko Inagawa, Mr. Peter Naglik, Ms. Tracy Theile, and Ms. Suzan Stevens, for their advice, thought-provoking discussion, and thoughtful support to this project from their specializations. I also wish to thank June Wilding Gilbert for her professional editing.

Lastly, I wish to express my heartfelt gratitude to my parents and sister for their loyal support.

## ABSTRACT

A challenge for English teachers working on EAP or IELTS preparation courses is to find methods to teach essay-writing more efficiently and effectively. The purpose of this study is to explore the role of model texts (sample answers) as a source of feedback in L2 writing. Using the IELTS writing test as focus material, the study focuses on what aspects of language ESL learners may notice by comparing their original texts to model texts in order to improve their writing skills. After classifying the learners' noticing into four categories (lexical, form, discourse and content), the study analyses the differences in quality and quantity of noticing, depending on (1) learners' proficiency and (2) the type of writing tasks.

The participants in the present study were 14 Japanese ESL learners (seven advanced learners and seven intermediate learners). After completing the two tasks in the IELTS writing test (descriptive and argumentative essays), the participants were asked to think aloud as they compared the two texts, followed by an interview to ascertain their general attitude towards learning L2 writing using model texts. The data classification procedure was partly replicated from the previous relevant L2 writing studies (Swain & Lapkin, 1995; Qi & Lapkin, 2001; Hanaoka, 2007) and the results were discussed based on the recent theoretical frameworks of attention, noticing, and awareness (Schmidt, 1990, 2001). The study reveals that there is a substantial difference in the quality and quantity of learners' noticing according to their proficiency and the type of tasks undertaken. Several pedagogical implications for L2 writing instructors and EAP course designers are also discussed.

## TABLE OF CONTENTS

	Page
STATEMENT OF AUTHORSHIP	i
ACKNOWLEDGEMENTS	ii
ABSTRACT	iii
TABLE OF CONTENTS	iv
LIST OF TABLES	vi
LIST OF FIGURES	vii
CHAPTER 1: INTRODUCTION	1
CHAPTER 2: IELTS WRITING TEST	4
2.1. General Information of IELTS	4
2.2. IELTS Writing Test	5
2.3. Assessment Criteria	5
2.4. Learning material for IELTS writing	6
CHAPTER 3: LITERATURE REVIEW	7
3.1. Output in SLA	7
3.2. Noticing in SLA	8
3.3. Noticing in L2 writing studies	10
3.4. Think-Aloud Protocol Approaches in L2 Writing Studies	12
3.5. Model Essay as a Feedback Tool	14
3.6. Research Questions	15
CHAPTER 4: METHODOLOGY	17
4.1. Participants	17
4.2. Data Collection Procedure	18

4.3. Instrument	20
4.4. Data Analysis	21
CHAPTER 5: RESULTS	26
5.1. Results for Research Question 1	26
5.1.1. General Tendency of the Frequency of LREs	26
5.1.2. Lexical LREs	29
5.1.3. Form LREs	30
5.1.4. Discourse LREs	31
5.1.5. Content LREs	32
5.1.6. Other LREs	33
5.2. Results for Research Question 2	33
5.3. Results for Research Question 3	37
CHAPTER 6: DISCUSSION	42
6.1. Data Analysis for Research Question 1	42
6.2. Data Analysis for Research Question 2	44
6.3. Data Analysis for Research Question 3	47
CHAPTER 7: CONCLUSION	49
7.1. Summary of the Study	49
7.2. Pedagogical Implications	49
7.3. Limitations of the Study and Suggestions for Further Research	51
REFERENCES	53
APPENDICES	59
Appendix A	59
Appendix B	61

**LIST OF TABLES**

	Page	
Table 1	Frequency, proportion, mean and standard deviation of the LREs	26
Table 2	The frequency and percentage of subcategories for each LREs	28
Table 3	Descriptive statistics of LREs by the advanced and intermediate groups	34
Table 4	Frequency of LREs at two levels of awareness	36
Table 5	Descriptive statistics of LREs between Task 1 and Task 2	37
Table 6	Numbers of students in each category of conceptions of learning and language learning across the three proficiency groups	39

**LIST OF FIGURES**

	Page
Figure 1 Frequency of the five LREs	27
Figure 2 Frequency of each LREs by the advanced and intermediate groups	34
Figure 3 Frequency of LREs at two levels of awareness	36
Figure 4 Frequency of each LRE between two tasks	38
Figure 5 Frequency of LREs at two levels of awareness	40

## CHAPTER 1

### Introduction

Over the decades, there have been a growing number of English learners wishing to study at tertiary level in English speaking countries. As a result, many English as Second Language (ESL) students are enrolled in English for Academic Purposes (EAP) courses which provide the opportunity to acquire essential skills for their prospective studies in English-medium universities (i.e., courses taught in English at universities in Anglophone or non-Anglophone countries). Besides acquiring academic skills, EAP courses have other aims, such as to support non-native English speaker students to go through English language tests such as the International English Language Testing Systems (IELTS), which is widely recognized as a language requirement for entering universities mainly in the United Kingdom, Australia, and New Zealand.

Academic essay-writing is one of the important skills which international students must acquire both for their prospective studies in English-mediated university courses and for obtaining the requiring score in the writing section of the IELTS. However, due to the complexities of learning to write well in a second language (L2), there has been vigorous debate on more efficient and effective teaching methods and feedback among many EAP teachers and L2 writing researchers.

A number of researchers admit that feedback plays an essential role in L2 writing instruction (e.g., Collins, Neville, & Bielaczyc, 2000; Ferris, 1995; Ferris & Hedgcock, 1998; Hedgcock & Lefkowitz, 1994; Hyland, 2003; Leki, 1990; Tickoo, 2001). There are several types of feedback in L2 writing instruction, such as teacher's essay correction, reformulation (native-speaker writing instructors' rewriting to keep the students' original ideas and correct only grammar mistakes and inappropriate

vocabulary), and peer feedback. A model essay written by native or native-like proficient writers may also be a beneficial resource if it can function as a feedback tool. The present study focuses on the role of model essays in the IELTS writing test and attempts to investigate whether model essays are a beneficial feedback tool or not.

Since 1970s, effectiveness of reformulation has been investigated by many researchers in the field of L2 writing (e.g., Cohen, 1987; Leki, 1990; Levenston, 1978; Thornbury, 1997). In recent L2 writing research, reformulation has been viewed as an insightful technique since this method has the advantage of urging students to notice language problems in their writing (Qi & Lapkin, 2001). However, some researchers point out that only practicing writing and accepting feedback are not sufficient to become a successful writer and L2 writing skills should be acquired by being exposed to various types of texts (Ferris and Hedgcock, 1998). From a practical viewpoint, reformulation can also be problematic as this technique is time-consuming for face-to-face instruction if the ratio of students to teacher is large.

This study attempts to investigate the role of a model essay as a feedback tool for L2 writing instruction. More specifically, it aims to reveal how Japanese ESL students notice their language problems by comparing their own writing with a model essay in the context of preparation for the IELTS writing test. The methodology is partly replicated from Qi and Lapkin's (2001) study on the reformulation method. By analyzing participants' verbal comments (think-aloud protocols) on what they noticed when comparing their own writing to the model essay, the tendency of their noticing was revealed. The results are expected to give a valuable insight into how EAP teachers could use model essays so as to improve students' academic writing skills. The three research questions explored in this study are as follows:

1. What aspects of language do Japanese ESL learners notice by comparing their own writing with sample essays?
2. Is there any difference in the noticing between more proficient learners and less proficient learners?
3. Is there any difference in the noticing between different types of writing tasks?

Following this introduction, Chapter 2 provides the general information of IELTS writing test with the details of assessment criteria. Chapter 3 gives an overview of the role of output and noticing in second language acquisition (SLA), the relevant empirical studies on L2 writing, the methodological issues on the think-aloud method, and the argument on the model essay approach in L2 writing. Chapter 4 describes the methodology of the present study. The results obtained are shown in Chapter 5, followed by the discussion of the data in Chapter 6. Finally, the summary of the study, pedagogical implications, the limitations of the study, and suggestions for further research are discussed in Chapter 7.

## CHAPTER 2

### IELTS Writing Test

#### 2.1. General Information of IELTS

IELTS is a set of examinations managed by the British Council, the University of Cambridge ESOL Examinations, and IDP IELTS Australia. IELTS has become increasingly recognized as a “secure, valid and reliable indicator of true to life ability to communicate in English for education, immigration and professional accreditation” (UCLES, 2007, p. 3). There are two modules (general training and academic) in IELTS and candidates can choose either of them according to their purpose for taking the test. The academic module is for candidates wishing to be enrolled in the degree courses of universities and other tertiary institutions, while the general training module is mainly for non-academic purposes such as immigration.

The test is divided into four sections: speaking, listening, reading, and writing. In the speaking and listening sections, the same tasks are distributed to both general and academic module candidates, while EAP students complete different reading and writing tasks from those for general training module candidates. The writing section consists of descriptive (Task 1) and argumentative (Task 2) essay-writing tasks. IELTS scores (academic module) have been used to determine whether the applicants have sufficient academic skills to follow studies at an undergraduate or postgraduate level in English speaking countries. IELTS has been adopted by many English-medium universities all over the world, mainly in the United Kingdom and Australia. Further, IELTS is now the test preferred by all universities in Australia (Coley, 1999).

## 2.2. IELTS writing test

In Task 1, examinees are given questions containing some visual information such as tables, graphs, charts, and diagrams. Examinees are expected to write a short description of information presented in the visual data. They must write at least 150 words in Task 1.

In Task 2, candidates are required to write an argumentative essay in response to a problem, opinion or controversial proposition. They have to show an ability to demonstrate an argument from a certain standpoint, suggest the solution, justify their opinion by drawing on their own knowledge and experience, weigh it against other opinions, and support their argument with their own experiences. The minimum word length is 250 in Task 2.

## 2.3. Assessment criteria

Candidates' writing scripts are assessed by trained and certified examiners. Each task is marked individually and the total score is reported on a scale of 1 to 9 at intervals of 0.5. Although the total band for the writing section is calculated by marking results for both, the weight of Task 2 is more than Task 1. The IELTS official assessment criteria for Task 1 and Task 2 scripts are the following (UCLES, 2007).

(1) Task Achievement (Task 1) / Task Response (Task 2): Task Achievement is the criteria on whether and/or to what extent candidates' writing successfully fulfilled task requirements using the limited number of words. Task Response is whether candidates construct an argument responding to the prompt given in the task and the argument is supported by their knowledge, personal experiences, solid evidence, and concrete examples.

(2) Coherence and Cohesion: Coherence and cohesion is the criterion regarding clarity and fluency of the passage. Specifically, coherence means how each sentence and paragraph are logically linked and cohesion is concerned with the proper use for cohesive device (e.g., conjunction, pronoun, repetition of synonym) to achieve the referential relationships between sentences and paragraphs.

(3) Lexical Resource: The criterion of lexical resource is concerned with the range of vocabulary the candidate uses and how appropriate the use of vocabulary is as a response to the prompt.

(4) Grammatical Range and Accuracy: This criterion examines how variedly and accurately the candidate uses grammatical resources at sentence-level.

#### 2.4. Learning material for IELTS writing

The most common material may be commercially produced textbooks designed for IELTS, which are available from many publishers. These textbooks differ according to the learning targets (from textbooks for comprehensive preparation for the IELTS to ones exclusively focused on certain skills). The writing section in these textbooks usually comprises exercise questions, useful resources of vocabulary and expressions, practice tests and the answer keys. Very often, answer keys refer to either ‘model essays (model answers)’ or ‘sample essays (sample answers).’ A model essay tends to refer to a model text written by a native or a writer of native-like proficiency, while a sample essay usually means a text composed by a non-native writer. In this study, a model essay composed by a native writer was used.

## CHAPTER 3

### Literature Review

In the field of L2 writing research, there has been much lively discussion with regard to the role of output and noticing. This chapter begins with a review of output and noticing in SLA (see 3.1 and 3.2), followed by previous L2 writing studies from which parts of the methodology of the current study were replicated (see 3.3) and the methodological problems of think-aloud approaches in L2 writing studies (see 3.4). Finally, the L2 researchers' attitude to the use of model essays in pedagogical contexts is discussed (see 3.5).

#### 3.1. Output in SLA

The role of input for SLA has been in debate and there have been various objections to Krashen's *Input Hypothesis*, which argues that L2 can be acquired if learners are exposed to comprehensible 'i+1' input. Swain (1985, 1998) investigated the learners in an immersion program in Canada and found that the students were lacking in accurate grammatical competence although they developed sufficient fluency to speak their L2. Thus, she concluded that *comprehensible input* is insufficient for acquiring L2 and *comprehensible output*, or output which is a little higher than the learner's present level, is necessary for obtaining accuracy and enhancing fluency in L2 learning.

The role of output in second language learning has been discussed mainly by Swain (1985, 1993, 1995, 1998, 2005) in her *Output Hypothesis*, which maintains that output has three major functions: noticing, hypothesis testing, and metalinguistic awareness. One of these, also known as the 'noticing/triggering function', has been claimed to play an important role for SLA. Swain (1995) claims that output promotes

noticing and “in producing the target language (TL), learners may encounter a linguistic problem leading them to notice what they do not know, or know only partially” (p. 129). In other words, output allows L2 learners to notice a gap existing between their interlanguage (IL) and TL, which may lead to learners’ conscious recognition of their language problems. Some empirical studies have demonstrated that output promotes noticing. Izumi’s (2002) study, for instance, demonstrates that visual input enhancement is not enough for acquiring the form of L2 learners’ IL; and that output facilitates noticing formal elements in the TL.

Testing hypotheses, in other words, trying out how the TL works is also an important function of output. Based on the assumption that output itself is the hypothesis, Swain (1995) states that “the output represents the learner’s best guess as to how something should be said or written” (p.132). As for metalinguistic function, she described that “under certain task conditions, learners will not only reveal their hypotheses, but reflect on them, using language to do so”. In addition to these three important functions of output, Swain (1995) also acknowledges ‘fluency’ as one of the important functions of output. De Bot (1996) discusses that the function of fluency is essential in that output promotes automaticity of the language processing system and enables language learners to “concentrate on a specific (sometimes form-related) aspect of language” (de Bot, 1996, p. 533). Among those four functions, the function of noticing is mainly discussed in the current study.

### 3.2. Noticing in SLA

In the recent SLA research, much emphasis has been placed on the role of attention, awareness, and noticing, which have been viewed as key issues in L2 learning

by many researchers (e.g., Ellis, 1993; Robinson, 1995; Schmidt, 1990, 1995, 2001; Schmidt & Frota, 1986; Swain, 1985, 1995; Swain & Lapkin, 1995). In his *Noticing Hypothesis*, Schmidt (1990) claims that noticing plays an essential role in SLA and that L2 learners must become 'aware' of certain aspects of language, mainly the meaning. Based on the assumption that awareness is significant for language learning, Schmidt emphasizes that awareness at the low level (noticing) is necessary and sufficient for SLA. Schmidt (1990) also states that learners need to notice all aspects of language equally, such as lexicon, grammatical form, sound, and pragmatic features. Some other researchers also claim that awareness is necessary for language learning. For instance, Robinson (1995) defines noticing as "detection plus rehearsal in short-term memory, prior to encoding in long-term memory" (p.296).

Schmidt (1995) distinguishes awareness at the low level (noticing) from awareness at the high level (understanding). He defines noticing as "conscious registration of the occurrence of some event" and understanding as "recognition of a general principle, rule or pattern". In other words, noticing deals with surface level language phenomena, while understanding is related to be learning at a more abstract level. Schmidt (2001) extends further the discussions regarding the role of noticing. Based on the assumption that noticing and understanding are different in the level of awareness and on the psychological view that attention is of limited capacity, he states that "limited attentional resources are directed first at those elements that carry message meaning, primarily lexicon, and only later, when the cost comes down, towards communicatively redundant formal features of language" (p.13). At the conclusion of his chapter, Schmidt emphasizes that what aspects of language an L2 learner notices in the input depends on the individual differences.

Even though there is general agreement on the importance of awareness and noticing, some disagreements also exist in the SLA literature. Tomlin and Villa (1994) disagree with Schmidt's *Noticing hypothesis* and point out the necessity of finer discussion about the concept of attention. Dividing the function of attention into alertness, orientation, and detection, they claim that detection, "the cognitive registration of sensory stimuli" (p. 192), which does not require conscious awareness, is vital for language learning. Robinson (1995) viewed noticing as "what is both detected and then further activated following the allocation of attentional resources" (p. 297), although he agrees with Schmidt in that awareness is necessary for SLA. Robinson (1995, 2003) particularly emphasizes that task complexity plays an important role by guiding learners' attention to certain aspects of language. Although there are a number of views on noticing, little is known about what aspects of language are noticed in L2 writing.

### 3.3. Noticing in L2 Writing Studies

The question about what aspects of language L2 writers notice was posed by Swain and Lapkin (1995). Their empirical study investigated the role of output in L2 writing context, examining whether the learners' output could allow them to become aware of language problems they encounter in composing. The participants, French-immersion students in Canada, were asked to speak (think aloud) whatever was on their mind in L2 composition. During analysis of think-aloud protocols, the units called 'language-related episodes (LREs)' were identified and categorized into several groups according to the type of language problems. The results revealed that the participants noticed language problems, which promoted them to modify their output.

Based on the results, they concluded that “noticing may occur because of either internal or external feedback which may prompt, for example, the generation of alternatives and assessment of them through simple inspection through to complex thinking (p.386).”

With a three-stage L2 writing task consisting of composing, comparing, and revising L2, Qi and Lapkin’s (2001) case study examined the role of noticing and reformulation as feedback. Two Chinese ESL students at different proficiency levels were asked to think aloud about whatever they noticed at the stage of comparing the learners’ own texts with the reformulated version of them. The think-aloud protocols were analyzed and categorized into lexical, form, and discourse LREs. The results indicate that there is a certain difference in the frequency of each LRE between more proficient and less proficient learners. The researchers also found that the more proficient student noticed with a higher level of awareness than the less proficient student. The researchers also assert that the reformulation technique allows learners to notice a gap between their IL and TL by comparing their own texts with the reformulated texts.

Based on Qi and Lapkin’s (2001) case study, Hanaoka (2007) investigated the role of model texts in promoting noticing in a four-stage study consisting of output, comparison, and two revision stages. In the comparison stage, the 37 participants, Japanese sophomore students at a woman’s university in Japan, were asked to write on whatever they noticed as they compared their original text with the models. For the sake of analysis, noticing was operationalized as self-reports in the form of note-taking. The data was coded into four categories; lexis, grammar, content, and other. The findings of this study indicate that the participants noticed the lexical aspects far more frequently than the other three categories. The results also reveal that more proficient learners

noticed language aspects significantly more frequently than less proficient counterparts. A further notable finding is that the participants noticed their respective linguistic problems and autonomously found solutions in the models.

Hanaoka's (2007) study adopted different methodologies from Qi and Lapkin (2001). In order to explore the characteristics of the participants' noticing, Hanaoka collected note-taking data with a one-paragraph short writing task (consisting of 70-80 words) while Qi and Lapkin adopted a think-aloud approach with a longer essay-writing task as an instrument. For the purpose of the present study, the IELTS writing test (multi-paragraph writing consisting more than 150/250 words) was used. The data correction procedure including the think-aloud stage was partially replicated from Qi and Lapkin (2001). However, there are several methodological problems regarding the think-aloud approach which L2 writing researchers should take into account.

#### 3.4. Think-Aloud Protocol Approaches in L2 Writing Studies

Some L2 writing studies have been conducted with think-aloud protocols as the data collection procedure. Among types of verbal reports, think-aloud protocols have several advantages. For instance, according to Smagorinsky (1989), think-aloud protocols can provide a rich amount of data and the obtained data reflect on the subject's mental processes. Schmidt (2001) argues that concurrent verbal reports such as think-aloud protocols are trustworthy evidence as to whether something has been consciously perceived or noticed. However, as Russo, Johnson and Stephens (1989) argue, think-aloud protocols can cause reactivity (the changes of mental processes due to thinking-aloud itself) and thinking aloud may eventually affect subjects' learning outcome. Ellis (2001) also doubts the validity of metalinguistic think-aloud approaches

where subjects perform some tasks and verbalize think-aloud protocols simultaneously. Thus, in order for L2 writing researchers to appropriately operationalize think-aloud protocols, these reactivity issues should be taken into consideration.

There has been little research on reactivity in SLA research, and even less in L2 writing studies. Leow and Morgan-Short (2004) examine the reactivity of concurrent nonmetalinguistic think-aloud protocols during the reading process. Although the results indicate that there is no reactivity, the researchers point out other potential factors which may cause non-reactivity, such as prior knowledge of the targeted items and the length of text. Bowles and Leow (2005) also investigate the reactivity of think-aloud protocols during L2 reading. Their results indicate that think-aloud does not yield reactivity although the experimental (think-aloud) groups spent more time completing tasks than the control group. They consider that several factors such as working memory, reading speed, and learning style, might affect reactivity.

Sachs and Polio (2007) is, as far as SLA research is concerned, the sole study that investigates the reactivity of think-aloud protocols using L2 writing tasks. Fifteen high-intermediate English ESL learners participated in a repeated-measures study (error correction, reformulation, and reformulation + think-aloud). They were asked to think aloud in L2. The results revealed that the 'reformulation' group significantly outperformed the 'reformulation + think-aloud' group. Although it appears that think-aloud may negatively affect the learners' performance, the researchers suggest that the potential factors of the negative reactivity might be the learners' insufficient L2 proficiency.

Given the limited number of studies, it is impossible to reach the conclusion that 'writing + thinking aloud' is reactive and that 'reading + thinking aloud' is not. In order

to reduce the effects of reactivity of the think-aloud method, the research design of the current study has no pre- and post-test design. In other words, the main focus of the present study is not on whether the use of model essays enhances learners' writing skills, but on what L2 learners can notice by comparing their original essays with a model essay. In addition, it might be the case that the reactivity for the current study is comparatively small since the comparing process is more 'reading + think-aloud' rather than 'writing + think-aloud'. In the next section, the argument about the role of model essays, as reading material for feedback, is discussed.

### 3.5. Model Essay as a Feedback Tool

Some L2 writing researchers argue that L2 learners should be encouraged to use a model essay for improving their writing skills in terms of the relationship between reading and writing. Ferris and Hedgcock (1998) argue that L2 writers have to be exposed to various types of reading material since it is difficult to acquire L2 writing skills by only writing. Eschholz (1980) points out that what L2 learners write depends on what they read and they can improve their L2 writing skills by reading. He also argues that given the opportunities to learn rhetorical modes, L2 learners can eventually apply their knowledge about those modes to their writing. Based on Cumming's (1995) empirical study, which demonstrates the significance of rhetorical aspects of texts in model essays, Smagorinsky (1992) discusses that model essays are the most helpful tool if L2 writers have a sufficient amount of content knowledge. Thus, some researchers emphasize the necessity of a model text illustrated in an academic writing textbook, which enables L2 writers to pay attention to the various aspects of TL (e.g., Hyland, 2003).

However, there are also several objections to using model essays in an L2 writing context. Murray (1980) points out that the process of making meaning in L2 cannot be achieved by referring to written texts. In addition, Goby (1997) asserted that model essays prevent L2 learners from having creativity, which she believes is one of the important aspects of L2 writing skills. Writing instruction with model essays has also been criticized by other researchers (Collins & Gentner, 1980; Judy, 1980) for laying emphasis not on content but on form. They insist that language form and the content of composition are inseparable. Even among researchers who claim that model essays can be beneficial pedagogical tools, there has been agreement that reading model essays is important but not totally sufficient (Ferris and Hedgcock, 1998; Hyland, 2003). However, there has been little empirical research to explore the role of model essays in L2 writing pedagogy.

### 3.6. Research Questions

Although there have been some L2 writing studies concerning the effectiveness of model essays (e.g., Hanaoka, 2007), no study has been conducted with academic essay-writing such as the IELTS writing test, as far as the author could find. As a partial replication of Qi and Lapkin (2001), the current study aimed to examine how L2 writers could receive beneficial feedback by using model essays. Pre- and post-tests, which Qi and Lapkin (2001) conducted to assess whether or not the participants successfully improved their writing skills by using model essays, were not carried out because the think-aloud method itself might affect their writing performance. Instead, the main focus of this study is on what aspects of language ESL Japanese learners notice in comparing their own essays and model essays. In order to explore the usefulness of

model essays of the IELTS writing test as a feedback tool, the focus of the current study was also on the difference in noticing caused by ‘learners’ proficiency’ (advanced or intermediate learners) and ‘task difference’ (Task 1 and Task 2). Thus, this study aims to answer the following research questions.

1. What aspects of language do Japanese ESL learners notice by comparing their own writing with model essays?
2. Is there any difference in the noticing between more proficient learners and less proficient learners?
3. Is there any difference in the noticing between different types of writing tasks?

## CHAPTER 4

### Methodology

#### 4.1. Participants

The participants in this study were 14 Japanese ESL learners, consisting of seven university students (two males and five females aged 21-35) and seven EAP students (one male and six females aged 21-32) enrolled in language schools in Brisbane, Australia. All 14 participants had studied English in Japan for more than six years at junior and high school levels.

Five of the university students were in their third semester of their degree programs and the remaining two university students were in the first semester (but both had enrolled in degree programs for more than a year in Canada and the United States) at the time of the data collection. The students' majors included education, international communication, Japanese translation and interpreting, and bio-technology. Besides the six years of learning experience, they had studied an additional five to seven years at university and language schools in Japan or Australia. All of them had taken the IELTS test before and their previous IELTS scores (bands) for the writing test ranged from 6 to 8.

The EAP students were enrolled in EAP courses. All of the EAP students had taken the IELTS test previously and achieved scores of 5 or below in the writing section. Besides learning experience at their junior and high schools, they had learned English for two to three years at university in Japan. At the time of the experiment, they had been attending English courses in Australia for between five and nine months.

In the present study, the participants were grouped either into the more proficient group or the less proficient group based on their learning experiences and IELTS scores.

In order to distinguish clearly the more proficient from the less proficient L2 writers, their essays were assessed by two native English-speaking teachers who were teaching EAP courses in Australia. The assessed essays were arranged in order from the best to worst and the two least proficient essays written by university students and the two best proficient essays written by EAP students were identified. After eliminating those four essays, the data of the five university students (the advanced group) and the other five EAP students (the intermediate group) were used for research question 2, which concerned the proficiency difference. The whole data (from all 14 students) was used for answering research questions 1 and 3.

#### 4.2. Data Collection Procedure

The data collection was conducted in a small and quiet room in the main library at The University of Queensland. To avoid misunderstanding, all the instructions were given in Japanese. The data collection procedure consisted of three stages.

Stage 1: Writing (in English for 60 minutes). Two types of writing tasks (see 4.3. Instrument) were given and each participant was then asked to write two essays within approximately one hour. As time management was not the main concern in this study, the participants were allowed to spend more than one hour, if necessary, to complete the tasks. However, they were *not* allowed to use dictionaries since one of the aims of the study was to investigate what the participants could notice just from their own texts and model essays, without any other aids.

Stage 2: Think-aloud (in Japanese; 30-40 minutes). At the beginning of this stage, each participant was trained to produce think-aloud protocols for a few minutes so that they could get accustomed to it and be free from anxiety and nervousness. After

the training, the sheets containing the model essays for Task 1 and Task 2 were given and each participant was asked to say whatever he/she noticed while comparing his/her own essays with the model essays. The verbal instruction which each participant received was the following: “Now, you are trying to improve your writing. You have a sample essay here. Please compare your own essay with the model essay and say whatever you notice. Any general or specific matters are okay.” Half of the participants were asked to compare with the Task 1 model first, then the Task 2 model. The remaining half of the participants were asked to do the opposite. The reason for this switch was that the researcher perceived that the participants tend to verbalize more in the second task than in the first task as they get used to think aloud during data collection.

The participants’ think-aloud protocols were recorded with an electronic recording device. The recording time ranged from 14 to 23 minutes for each comparison. If a participant stopped verbalizing within 10 minutes, the researcher urged “Could you go over both texts checking if there is anything else to notice?” In order to equalize the opportunity to receive prompts from the researcher, participants who verbalized for more than 10 minutes were also urged in the same way.

Stage 3: Interviews (in Japanese; 10 minutes). After the think-aloud session, the researcher interviewed each participant to gain further insight into his/her attitude to using model essays to improve his/her writing skills. The interview comprised three questions as follows:

- (1) Do you think that using model essays is helpful for improving your writing? Why?
- (2) Which model essay is the more beneficial one for you, Task 1 or Task 2? Why?

- (3) Besides model essays, what do you need in order to obtain more beneficial feedback? Please describe what, if anything, this model essay approach lacks.

After the three stages, the think-aloud protocols and the comments in the interviews were transcribed by the researchers.

#### 4.3. Instrument

The writing task given to each participant is an exercise version of IELTS writing test (academic module) in a commercially published IELTS preparation textbook edited by Scovell, Pastellas, and Knobel (2004) (see Appendix A). The writing test includes two writing tasks: descriptive and argumentative essays. There is a minimum requirement for word length (at least 150 words for Task 1, and at least 250 words for Task 2).

Task 1 is a descriptive report of a table about tourism in the United Kingdom. The task prompt is “Tourism in the United Kingdom contributes billions of pounds sterling to the UK economy. The table below identifies the twelve most visited paid-admission attractions in the UK in 1999 and 2000. Write a report describing the information shown in the table.” Due to the nature of this essay (in descriptive essays, it is not necessary for each writer to show their opinions, experiences, and knowledge), what students composed tended to be similar to the model essay.

Task 2 was an argumentative essay on the topic of ‘capital punishment’, which was reasonably familiar to all the participants as they had previously discussed this issue in either academic circumstances, or in their daily lives (e.g., when watching TV programs). In order to avoid the consequence where the students’ essays were totally

different from the model essay's argument, a yes/no question was not selected as material. Therefore, the prompt for Task 2 was "Discuss when, if ever, capital punishment can be viewed as a valid punishment for crime." None of the participants had ever written essays on these topics before.

The model essay employed in this study was also from the same textbook (see Appendix B). The two native English-speaking teachers admitted that the model is at native-writer level and substantially more advanced than the best writer of all 14 participants.

#### 4.4. Data Analysis

The data analysis methods were partly replicated from Qi and Lapkin's (2001) study, where noticing was operationalized as the participants' verbalized language-related episodes (LREs), which indicates that the number of LREs is interpreted as the frequency of noticing. First, the students' recorded comments were transcribed into think-aloud protocols. Each student's think-aloud protocols were then divided into segments (LRE) by identifying the interval between one episode and another, and the gap between the meaning of each episode.

After that, LREs were classified into three broad categories, lexical, form, and discourse. In the current study, based on Qi and Lapkin's (2001) classification, the LREs were categorized as following: (1) lexical – selecting words, phrases, and expressions; (2) form – articles, plural, sentence structure, verb form, tense, prepositions, comparative and superlative, punctuation, and spelling; (3) discourse – logical sequencing (cohesion and coherence), organization of paragraphs, inter-sentential relationship, and cohesive devices.

However, the present data revealed that students verbalized a number of episodes which were related to the content of the texts. Thus, in this study, the researcher added one more category; (4) content – which information in the table should be included in the essay (Task 1); own opinions, knowledge, experiences, evidences, counter-evidences, and supportive ideas (Task 2). Although content LREs might not have been precisely ‘language-related’, they were named ‘content LREs’ in this study for the sake of convenience. The LREs which did not fall into any group above were categorized into (5) other LREs. The following examples present briefly the way each LRE was coded into four major categories. In all the examples, the italics indicate the part verbalized by each participant in L1 (Japanese). L2 (English) use was presented by using underline.

(1) “*This word, ‘Clearly’ is used in the model essay. I didn’t use this type of word in my essay.*” ( **Lexical** LRE – from Task 1 – Student 2)

(2) “*In my writing, in line 11, I used present tense, actually I had to use the past tense, because this information is about in 1999.*” (**Form** LRE – from Task 1 – Student 3)

In the first example, Student 2 mentioned the usage of a particular word “clearly” and he noticed that he did not use the same type of vocabulary in his essay. Student 3 found the misuse of tense which is related to accuracy of form, as a result of comparison in the second example. She noticed that the right tense was not present tense but past tense.

- (3) *“the model essay doesn’t use any words to connect sentences. But it’s logical enough because the reason comes just after the claim. (Discourse LRE – from Task 2 – Student 6)*

The LRE in example 3 is with regard to the discourse matter (inter-sentential sequence) which Student 6 encountered in the model essay. He described how the model essay achieved the logical inter-sentential sequence between two sentences.

- (4) *“In the third paragraph, the model essay described the numbers of visitors in detail. But I didn’t.” (Content LRE – from Task 1 – Student 14)*

Example 4 is related to the content. Student 14 compared the model essay to her own text and found that the model contained certain information (the number of visitors) but her essay did not.

These procedures were repeated by a Japanese researcher who had conducted several SLA studies and was familiar with the relevant knowledge. First, the two researchers (the author and the researcher) discussed the categorization criteria and then modified them. After reaching agreement on the categorization criteria, both the author and the researcher then analyzed four participants’ think-aloud protocols and coded the data into the five categories above. The intercoder reliability for categorization of LREs was 90 %. After further discussion, an agreement on categorization of LREs was achieved by two researchers.

After the data collection, it was found that the students’ noticing behaviour was

different according to whether or not the information included in each LRE was detailed. In order to explore the quality of noticing identified from the LREs, two more categories were borrowed from Qi and Lapkin (2001). In their study, each LRE was categorized into (1) noticing only (the case where a participant just notices a certain aspects of texts without giving any reasons) or (2) accept a model with a reason (the case where participants accepted the superiority of the model by pointing out specific reasons). The researchers named the former noticing “perfunctory” and the latter “substantive”, and these two categories were interpreted as the difference between awareness at a low level (noticing only) and high level (accepting model essays with a reason). The following are examples of LRE which were categorized into ‘noticing only’ and ‘accepting model essays with a reason’, respectively.

(5) “*The model says ‘the most significant decrease was experience’. The word ‘experience’ is good. I can’t use this word.*” (Lexical – **noticing only** – from Task 2 – Student 1)

(6) “*Here, the model uses ‘offender’. That is good expression. It’s very appropriate word in this type of essay.*” (Lexical – **accepting model essays with a reason** – from Task 2 – Student 12).

In both example 5 and 6, the focus of the LREs is a lexical aspect. In example 5, however, Student 1 just mentioned the lexical items she noticed. As she did not describe any specific reason, this LRE was categorized into ‘noticing only’. On the other hand, Student 12 paid attention to the particular word in the model and described the specific

reason why the word ‘offender’ was appropriate. Thus, example 6 was categorized into ‘accepting model essays with a reason’. Likewise, examples 1 to 4 can be categorized in the same way.

After all the data were categorized, the frequency of the major five categories (lexical, form, discourse, content, other) and subcategories of LREs were identified and the descriptive statistics for each category were calculated so as to answer the first research question. In order to response to the second research question, the descriptive statistics of the LREs were compared between the two proficiency groups (advanced *vs.* intermediate). Finally, the data were compared between the two different writing tasks (Task 1 *vs.* Task 2)

## CHAPTER 5

### Results

#### 5.1. Results for Research Question 1

The first research question aimed to investigate which aspects of language the participants noticed. In order to explore what aspect of language they were paying attention to, the subcategories for each LRE were identified. In 5.1.1, the frequency of five main LREs (lexical, form, discourse, content, and other) which were verbalized by all the participants and of their subcategories are quantitatively analyzed. The subcategories are described with some examples from 5.1.2 to 5.1.6.

##### 5.1.1. General Tendency of the Frequency of LREs

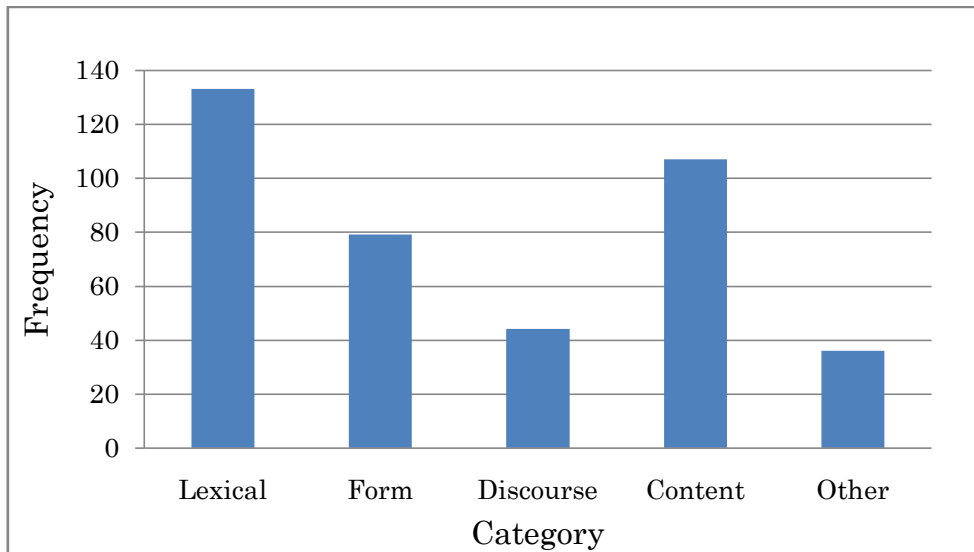
Table 1 summarizes the frequencies, proportions, mean, and standard deviation of the LREs verbalized by all participants (frequencies are also visualized in Figure 1).

Table 1

*Frequency, proportion, mean and standard deviation of the LREs*

	All participants ( $N = 14$ )			
	<i>n</i>	<i>%</i>	<i>M</i>	<i>SD</i>
Lexical	133	33.3	9.50	4.15
Form	79	19.8	5.64	3.50
Discourse	44	11.0	3.14	1.92
Content	107	26.8	7.64	2.92
Other	36	9.0	2.57	1.34
Total	399	100.0	28.50	5.73

Figure 1

*Frequency of the five LREs*

As the table shows, the participants verbalized 28.50 LREs on average. The total number of LREs per participant ranged from 23 to 39. The largest proportion and mean frequency of the LREs was lexical (33.3%,  $M = 9.50$ ), followed by content (26.8%,  $M = 7.64$ ), form (19.8%,  $M = 5.64$ ), discourse (11.0%,  $M = 3.14$ ), and other (9.9%,  $M = 2.57$ ).

Several subcategories were identified in the three LREs (lexical, form, and discourse). Table 2 summarizes the cumulative frequency and the percentage of each subcategory. As can be seen from the Table 2, the frequencies of each subcategory were widely varied.

Table 2

*The frequency and percentage of subcategories for each LREs*

Category/Subcategory	Frequency	Percentage
<u>Lexical</u>		
Vocabulary and expressions	99	77.4
Lexical variety	21	15.8
Other	13	9.8
Subtotal	(133)	(100)
<u>Form</u>		
Grammar	49	62.0
Punctuation	18	22.8
Spelling	5	6.3
Other	7	8.9
Subtotal	(79)	(100)
<u>Discourse</u>		
Paragraph	21	47.7
Lexical cohesion	14	31.8
Other	9	20.5
Subtotal	(44)	(100)
<u>Content</u>	107	
<u>Other</u>	36	
<u>Total</u>	399	

It is obvious from the table that not only which aspects of language the students noticed but also how they used the model essays was widely varied. The following sections serve to illustrate the individual characteristics of subcategories for the five main LREs verbalized by the students.

### 5.1.2. Lexical LREs

The lexical LREs were the most frequently verbalized by the participants (133 LREs). The majority of the lexical LREs referred to the discovery of more appropriate, sophisticated, and academic vocabulary in the model essay (99 LREs, 77.4% of all the lexical LREs). The participants often accepted the superiority of the model texts in terms of vocabulary with such comments as, for instance, “I should have used this,” “This is much better,” or “I knew these words, but I couldn’t use them.” The following examples are common lexical LREs.

- (1) *“The model uses ‘respectively’. Yeah, this word is useful. I also used ‘respectively’ quite often. I think I used ‘respectively’ three times” (Student 6)*
- (2) *“I should have used ‘guilty’ or ‘innocent’, more suitable vocabulary for the topic of crime and law.” (Student 9)*

The participants’ attention was not necessarily drawn to particular lexical items. The 13 (out of 14) participants noticed the necessity of using more varied words and expressions at least once (21 LREs, 15.8%). Student 5 reported, for example:

- (3) *“This word ‘valid’ is good. I just kept using ‘important’. Maybe I should use more varied words” (Student 5)*

Some specific lexical items such as the use of an adverb, ‘respectively’ (7 participants noticed) and a phrase, ‘followed by’ (5 participants noticed) were found to

be verbalized more frequently than other lexical items. Although some LREs referred to ‘cohesive use of vocabulary’, this type of LREs were categorized into discourse LREs in this study (see 5.1.3).

### 5.1.3. Form LREs

With regard to form LREs (79 LREs), the participants tended to pay more attention to ‘accuracy’ than ‘usefulness’ or ‘variety’. Thus, the form LREs were often verbalized with the utterances such as “This is a mistake...,” “Good, I used right form,” or “Is it really correct?” Although some participants correctly solved the grammatical problems by referring to the model essay, others failed to do so even though they carefully compared their own texts with the model. The following are the examples of form LREs.

(4) “*The model says ‘beyond reasonable doubt’. I would say ‘doubts’. I have to use ‘-s’ correctly.*” (Student 1)

(5) “*I used present tense. Do I really have to use past tense for this task?*” (Student 7)

The most frequently commented on form LREs concerned grammatical problems (49 LREs, 62.0 % of all the form LREs), followed by punctuation (18 LREs, 22.8%), and spelling (5 LREs, 6.3%). The common grammatical problems which the students reported were concerned mainly with tense (like example 5 above), verb endings, articles, prepositions, and auxiliary verbs. With regard to punctuation, 12 students out of 14 commented on the use of punctuation such as comma, colon,

semi-colon, quotation mark, and parentheses as follows:

- (6) *“I think the model uses comma in the right places. For example, here, it says ‘in the essay’ plus comma. The comma made this sentence easy to read.”* (Student 5)

In comparison, although it was difficult for some students to solve the grammatical problems, they tend to successfully find correct usage of punctuation and correct spellings by comparing their own essay with the model essay. The example of spelling LRE is as followed:

- (7) *“Oh! The spelling of ‘visiter’ is wrong. This should be ‘visitor’. I didn’t notice it until I read the model”* (Student 10)

#### 5.1.4. Discourse LREs

The discourse LREs, the least frequently verbalized among the four major categories (44 LREs), referred mainly to the organization of paragraphs and inter-paragraph cohesion (21 LREs from 12 participants). See the following example.

- (8) *“First of all, my writing is not clearly structured. It doesn’t have any of introduction, body, and conclusion like the model essay”.* (Student 12)

A number of the participants paid attention to ‘lexical cohesion’ as well as inter-paragraph and inter-sentence cohesion. Eleven participants verbalized 14 LREs pertaining to paraphrased or repeated words/expressions especially when they were

referring to the model essay, as follows:

- (9) *“The model used different words in different places. Here, it used ‘capital punishment’ but, here is ‘death penalty’. That makes these sentences look more sophisticated.”* (Student 4)

Although the border between lexical and discourse LREs tended to be vague, in this study, the LREs concerning lexical problems were classified into either lexical or discourse categories, according to which aspect was predominant. For instance, example 9 above was categorized into discourse LRE (lexical cohesion) since Student 4 seemed to be focusing mainly on paraphrasing rather than just lexical variety.

#### 5.1.5. Content LREs

The content LREs were the second most frequently verbalized by the participants (107 LREs). No sub-categories were identified within this category. However, there were some differences in this category according to which aspects of content the students were focusing on, especially in Task 2 (e.g., supportive/counter evidence, own experience, examples, and general knowledge). On the other hand, in Task 1, all of the verbalized content LREs were concerned with what information in the table should be included in their descriptive essay. The following are the examples of the content LREs in Task 1 and Task 2, respectively.

- (10) *“In my essay, I mentioned the most popular and the least popular attractions. Basically, the model says the same but it includes more detailed information about*

*the number of visitors.”* (Student 6; in Task 1)

(11) *“The way the model contains a lot of examples in each paragraph. That makes the whole argument persuasive.”* (Student 5; in Task 2)

#### 5.1.6. Other LREs

Thirty-six LREs were not categorized into any of the four major groups above in the study. The majority of other LREs are concerned with the word number, spacing between paragraphs, abbreviations, and notation of numeric information.

#### 5.2. Results for Research Question 2

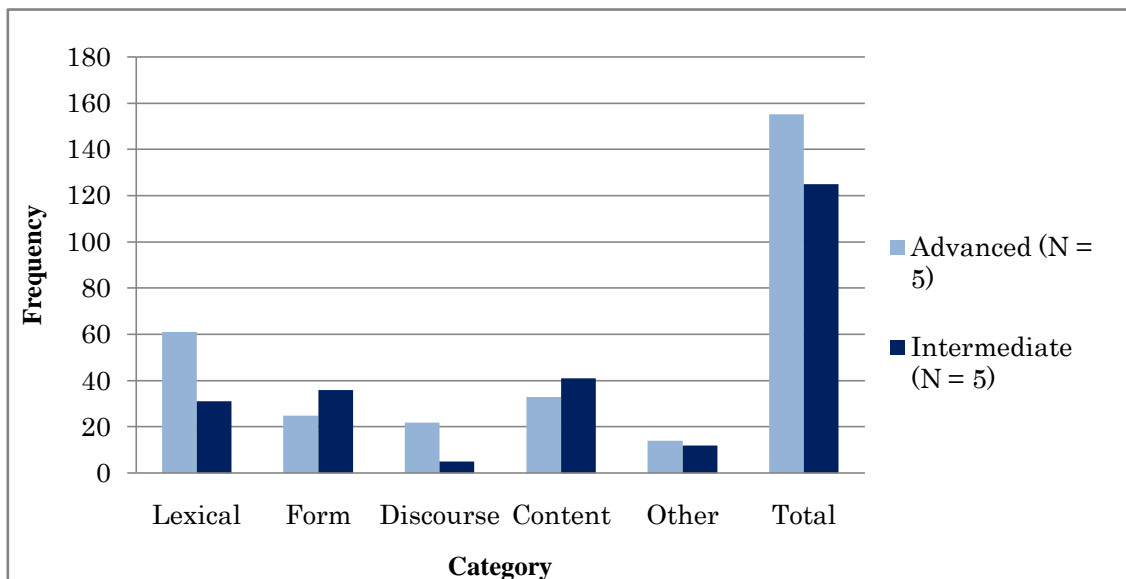
The second research question concerned the difference in the frequency of noticing between more proficient learners and less proficient learners. In order to answer the question, the data from the 10 participants which consisted of five most-proficient university students (the advanced group) and five least-proficient EAP students (the intermediate group) were quantitatively analyzed. Table 3 shows frequencies, proportions, mean, and standard deviation of the LREs produced by the advanced and intermediate groups (frequencies are also visualized in Figure 2). As can be seen in the table, the total number and mean frequency of the LREs produced by the advanced group were substantially larger than the intermediate group (155 LREs,  $M = 31.00$  vs. 125 LREs,  $M = 25.00$ ). However, some notable differences in the frequencies and proportions among categories can be detected between the two groups.

Table 3

*Descriptive statistics of LREs by the advanced and intermediate groups*

	Advanced (N = 5)				Intermediate (N = 5)			
	<i>n</i>	%	<i>M</i>	<i>SD</i>	<i>n</i>	%	<i>M</i>	<i>SD</i>
Lexical	61	39.4	12.20	3.49	31	24.8	6.20	3.11
Form	25	16.1	5.00	3.32	36	28.8	7.20	4.32
Discourse	22	14.2	4.40	0.89	5	4.0	1.00	1.22
Content	33	21.3	6.60	2.07	41	32.8	8.20	3.96
Other	14	9.0	2.80	1.92	12	9.6	2.40	0.89
Total	155	100.0	31.00	6.96	125	100.0	25.00	1.41

Figure 2

*Frequency of each LREs by the advanced and intermediate groups*

The advanced group reported nearly 40% of the lexical LREs of all categories, followed by content (21.3%), form (16.1%), discourse (14.2%), and other (9.0%), respectively. The most frequently verbalized LREs by the intermediate group are content LREs (32.8%), followed by form (28.8%) and lexical (24.8%). However, the

most distinctive feature here is the obvious lack of discourse LREs (4.0 %). In comparison to the intermediate group, the advanced group verbalized the higher percentage of the lexical (39.4%,  $M = 12.20$  vs. 24.8%,  $M = 6.20$ ) and discourse (14.2%,  $M = 4.40$  vs. 4.0%,  $M = 1.00$ ). However, the higher percentage of LREs was produced by the intermediate group in terms of form (28.8%,  $M = 7.20$  vs. 16.1%,  $M = 5.00$ ) and content (32.8%,  $M = 8.20$  vs. 21.3%,  $M = 6.60$ ).

Table 4 and Figure 3 display the frequency of the three main LREs (lexical, form, and discourse) between the advanced and intermediate groups, at different levels of awareness (noticing only vs. accepting a model with a reason). As can be seen in the table, the advanced group produced a larger number of LREs which were categorized into ‘noticing only’ (58 LREs, 53.7%) than ‘accepting model with a reason’ (50 LREs, 46.3%), while a much larger number of LREs verbalized by the intermediate group were categorized into ‘noticing only’ (58 LREs, 80.6%) compared with ‘accepting model with a reason’ (14 LREs, 19.4%). In addition to the overall inter-group difference observed, varied tendencies can be detected among three major categories.

For example, concerning form and discourse LREs, the advanced group accepted the model with a reason (14 form and 13 discourse LREs) slightly more frequently than they did with no reason (11 form and 9 discourse LREs), whereas the opposite case is true with regard to lexical LREs, which were more frequently verbalized by only noticing without giving a reason (38) than by accepting the model with a reason (23). On the other hand, the intermediate group rarely accepted the model essay with a reason (3 lexical and no discourse) compared with the number of LREs verbalized by noticing without reasons (28 lexical and 5 discourse). However, a comparatively larger number of form LREs were verbalized with a reason (11 out of 25)

among the lexical and discourse LREs.

Table 4

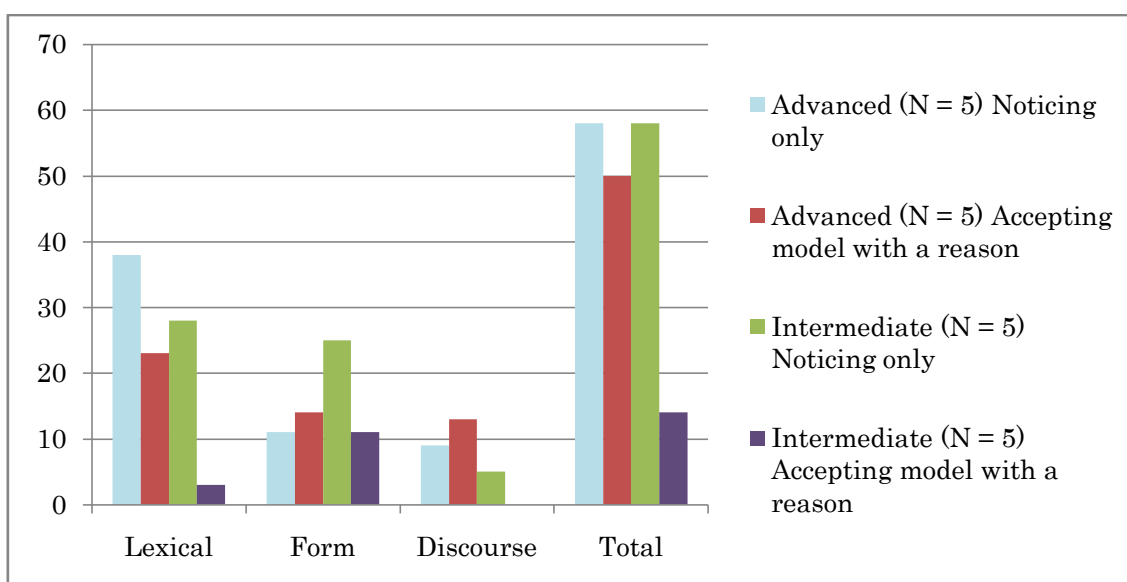
*Frequency of LREs at two levels of awareness*

	Advanced ( $N = 5$ )			Intermediate ( $N = 5$ )		
	Noticing only	Accepting model with a reason	Total	Noticing only	Accepting model with a reason	Total
Lexical	38	23	61	28	3	31
Form	11	14	25	25	11	36
Discourse	9	13	22	5	0	5
Total	58 (53.7)	50 (46.3)	108 (100)	58 (80.6)	14 (19.4)	72 (100)

*Notes.* Numbers in parentheses indicate the percentage.

Figure 3

*Frequency of LREs at two levels of awareness*



The second research question explored how the proficiency difference affected the tendency of noticing in comparison of the two texts. The aspects of language the students noticed differed between the advanced and intermediate group. Moreover, it turned out that the difference in noticing behaviour between ‘noticing only’ and ‘accepting model with a reason’ was also substantial between the two groups.

### 5.3. Results for Research Question 3

The third research question was concerned with comparing the task effects between Task 1, descriptive essay-writing, and Task 2, argumentative essay-writing. In order to answer the question, the data for all 14 participants were quantitatively analyzed. Descriptive statistics of the data were calculated to examine whether there was any difference in the frequency of LREs originating between Task 1 and Task 2. The results are summarized in Table 5 and graphically presented in Figure 4.

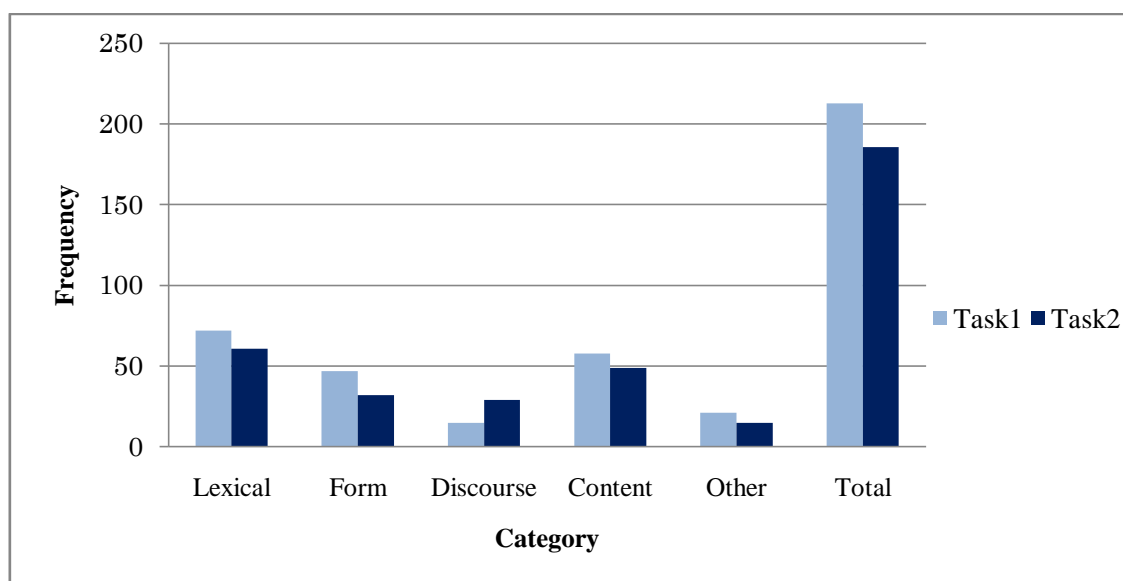
Table 5

*Descriptive statistics of LREs between Task 1 and Task 2*

	Task1				Task2			
	<i>n</i>	<i>%</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>%</i>	<i>M</i>	<i>SD</i>
Lexical	72	33.8	5.14	2.80	61	32.8	4.36	2.65
Form	47	22.1	3.36	2.41	32	17.2	2.29	1.49
Discourse	15	7.0	1.07	0.73	29	15.6	2.07	1.38
Content	58	27.2	4.14	1.46	49	26.3	3.50	2.24
Other	21	9.9	1.50	0.94	15	8.1	1.07	0.83
Total	213	100.0	15.21	3.21	186	100.0	13.29	3.91

Figure 4

*Frequency of each LRE between two tasks*



Total numbers of LREs which the 14 participants produced in Task 1 and Task 2 were 213 and 186, respectively. Taking the word length (150 words for Task 1; 250 words for Task 2) into consideration, it is probable that the number of LREs produced by Task 1 might be more than 213 if the word number for both tasks is the same. Therefore, one could say that Task 1's model is arguably more resourceful than Task 2's model in terms of noticing.

Compared with the case of Task 2, the larger number and proportion of lexical (33.8 %,  $M = 5.14$  vs. 32.8%,  $M = 4.36$ ), form (22.1%,  $M = 3.64$  vs. 17.2%,  $M = 2.29$ ) and content LREs (27.2%,  $M = 4.14$  vs. 26.3%,  $M = 3.50$ ) were derived from Task 1's model essay. However, the number and percentage of counted discourse LREs for Task 1 was nearly half of the ones for Task 2 (7.0%,  $M = 1.07$  vs. 15.6%,  $M = 2.07$ ).

Table 6 and Figure 5 illustrate the frequency of the two different types of noticing behaviour (noticing only and accepting model with a reason) between the two

different writing tasks. As can be seen from the table, there was not a substantial difference in the ratio of ‘noticing only’ to ‘accepting model with a reason’ between Task 1 (66.4% : 33.6%) and Task 2 (67.2% : 32.8%). Thus, there seems to be no task differences between Task 1 and Task 2 in terms of the level at which the students notice the three major aspects of language. However, this 2:1 ratio of the two types of noticing behaviour could not explain all the ratios across the three LRE categories (lexical, form, and discourse).

Table 6

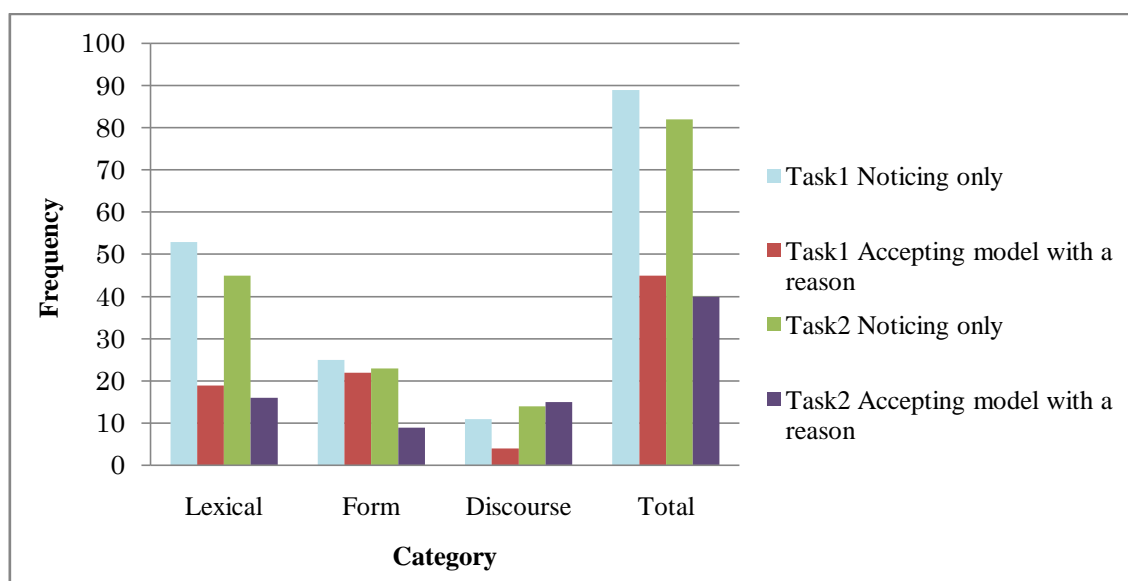
*Frequency of LREs at two levels of awareness*

	Task1			Task2		
	Noticing only	Accepting model with a reason	Total	Noticing only	Accepting model with a reason	Total
Lexical	53	19	72	45	16	61
Form	25	22	47	23	9	32
Discourse	11	4	15	14	15	29
Total	89	45	134	82	40	122
	(66.4)	(33.6)	(100)	(67.2)	(32.8)	(100)

*Notes.* Numbers in parentheses indicate the percentage.

Figure 5

*Frequency of LREs at two levels of awareness*



In task 1, the number of lexical LREs which were categorized into ‘noticing only’ (53 LREs) was considerably larger than ‘accepting model with a reason’ (19). The number of discourse LREs also showed the similar tendency between the two types of noticing behaviour (11 for ‘noticing only’ and 4 for ‘with a reason’). On the contrary, there was not substantial difference in number of form LREs between the two categories, ‘noticing only’ (22) and ‘accepting model with a reason’ (25).

In task 2, the number of both lexical (45) and form LREs (23) coded into ‘noticing only’ were substantially larger than both LREs (16 for lexical and 9 for form LREs) categorized into ‘accepting model with a reason’. However, a similar number of discourse LREs was produced between the two types of noticing behaviour (14 for ‘noticing only’ and 15 for ‘with a reason’).

The third research question investigated task differences in noticing aspects of language in comparison of the original essay and the model essay. As a whole, the

number of LREs originated in Task 1 was larger than Task 2 except for discourse LREs, which were produced less frequently in Task 1 than Task 2. Some substantial differences in number of LREs were also identified between students' noticing behaviour ('noticing only' or 'accepting model with a reason') in both writing tasks.

## CHAPTER 6

### Discussion

#### 6.1. Data analysis for Research Question 1: *What aspects of language do Japanese ESL learners notice by comparing their own writing with sample essays?*

The results of the first research question revealed a general tendency among Japanese ESL learners towards noticing language aspects in comparing their own writing with model essays. Broadly, a substantial number ( $M = 28.50$ ) of LREs was verbalized by each participant, which indicated that model essays written by native or native-like proficient writers could be a useful resource for encouraging L2 writers to notice various aspects of language.

In general, the participants noticed lexical aspects most frequently among the four major categories. This finding is consistent with some of the previous L2 writing studies which include a comparison-stage of students' original writing and its reformulated version (Swain & Lapkin, 1995, 2002, Qi & Lapkin, 2001) and a comparison of students' original writing and a model text (Hanaoka, 2007). Hanaoka's (2007) results show that noticing of lexical aspects occurred far more frequently (63.0%) than noticing of aspects of grammar (4.3%) and content (29.0%). The results in the current study, however, reveal a smaller gap among the proportions for these categories: lexical (33.3%), form (19.8%) and content (26.8%). With regard to writing tasks, Hanaoka (2007) adopted a writing task consisting of just one paragraph, whereas longer essay-writing tasks containing more than two paragraphs were implemented in the present study. Therefore, it could be argued that the longer an essay is, the more various aspects of language a learner notices.

The proportion of frequencies among the three major (purely 'language-related')

LREs (lexical, form, and discourse) verbalized in the current study was similar to that in Qi and Lapkin's (2001) case study, which explored the effectiveness of reformulation as feedback for L2 writing. Their results revealed that lexical aspects were the most frequently commented on, followed by form and discourse. Thus, it seems that the results of both studies accord with each other in terms of the proportion of LREs. Comparing the present study and Qi and Lapkin (2001), it can be argued that the model essays can provide opportunities to encourage learners' noticing of language aspects, which are similar to the ones that reformulated learners' texts could provide.

The results also reveal that there are both more common and less common subcategories for each LRE. For instance, the most common discourse LREs were related to paragraphs. In Qi and Lapkin (2001), participants verbalized little on the issue of paragraphs. It could be speculated that the reformulated texts did not contain any feedback about the use of paragraphs since reformulation can only correct grammatical or lexical mistakes but not the writer's original ideas. In Hanaoka (2007), the same situation is observed. However, in his study, there is no discourse category which falls into language-related episodes (LREs) since one-paragraph short essay-writing tasks were adopted. These findings infer that model essays, exemplifying the longer essay-writing tasks of the IELTS writing test, would be beneficial for leading L2 learners to pay more attention to paragraphs. On the other hand, reformulation and model texts for short writing tasks may not be particularly useful for encouraging L2 learners to notice discourse in L2 composition.

Curiously, it seems that the students verbalized each LRE in different ways. One of the fundamental differences between lexical and form LREs is that the students tended to search for more appropriate vocabulary, while they attempted to correct their

grammar accurately. Although the majority of the students were satisfied with the quality of the model essay as a feedback tool, some students were not. In response to the first question in the interview (“Do you think that using model essays is helpful for you to improve your writing? Why?”), two EAP students answered negatively. They found it difficult to solve grammatical problems, such as tense, since they failed to obtain the feedback on which tense was correct (or more appropriate). Furthermore, in response to the third question (“Besides model essays, what do you need in order to obtain more beneficial feedback? Please describe what, if anything, this model essay approach lacks.”), all the five EAP students in the intermediate group and two university students emphasized the necessity of native English teachers’ written corrections, particularly demanding the correction of grammatical mistakes.

In the present study, the content LREs are the second most frequently reported. It could be interpreted that a comparison of their own writing and model essays promotes learners to pay much attention to content, since model essays can provide completely novel perspectives on fulfilling the requirements of writing tasks. This could be difficult to achieve by using the reformulation approach, which basically maintains content of essays.

## 6.2. Data analysis for Research Question 2: *Is there any difference in the noticing between more proficient learners and less proficient learners?*

The results for the second research question, which are concerned with the effects of proficiency on the frequency of each LRE, revealed that there were some distinctive differences in the quantity and quality of noticing between the advanced and intermediate groups of Japanese ESL learners. Overall, the fact that the advanced

students verbalized more LREs than the intermediate group may confirm Schmidt and Frota's (1986) observation that "those who notice most, learn most" (p. 313), and Schmidt's (2001) discussion that "since beginning learners are cognitively overloaded, they cannot pay attention to all meaningful differences at once" (p. 7).

Astika (1993) reports that vocabulary is an important factor in accounting for the L2 writing performance. Astika's study reveals that 83.75% of proficiency of L2 writing can be predicted from vocabulary. The finding in the present study, namely that the advanced students tended to pay more attention to lexical aspects than the other aspects (form, discourse, and content), seems to be consistent with the results of not only Astika (1993) but also Qi and Lapkin's (2001) and Hanaoka's (2007) studies. This suggests that the more proficient learners tend to notice more lexical aspects of language in comparing of the two texts regardless of whether the feedback tool is reformulation or model essays.

On the other hand, the intermediate learners pay more attention to aspects of language form than the advanced learners. In addition, they pay little attention to discourse. These two facts also seem to be consistent with the results in Qi and Lapkin's study. That might suggest that the less proficient L2 writers notice more grammatical aspects but tend to fail to notice discourse aspects in comparison of the two texts, regardless of types of feedback (reformulation or model essays). According to Scarcella and Oxford (1992), some L2 writers tend to pay attention to the accuracy of each sentence, spelling, and punctuation rather than the other aspects of their writing. Porte (1996) also points out that some L2 writers are inclined to focus on each word but not to be involved in paragraph-level modification. Paradoxically, these points of view may be interpreted to mean that less proficient writers, who tend to fail to notice inter-sentential

relationships and paragraphs (discourse) tend to notice general grammar, punctuation, and spelling (form) more frequently than more proficient writers. Some of the Japanese ESL students' inclination to lay too much emphasis on language form might be partly due to the nature of the Japanese education system, which overemphasizes the importance of grammar instruction.

The results of the current study also revealed that the students with a higher level of L2 proficiency tended more frequently to verbalize specific reasons why they accepted the superiority of the model. The same result can be found in Qi and Lapkin's (2001) study, where the more proficient writer accepted the reformulation with a specific reason, whereas the less proficient writer rarely did so. Based on the argument of Qi and Lapkin, the finding in the present study that the advanced group produced more 'with-a-reason' LREs than the intermediate group can be interpreted to mean that the advanced group more frequently noticed the gap between IL and TL with a high level of awareness (understanding). On the other hand, the fact that the intermediate group often verbalized 'noticing-only' LREs may indicate that they noticed the gap with a lower level of awareness (noticing). These views accord to the words of Qi and Lapkin that "learners with a lower level of L2 proficiency may have more difficulty identifying the nature of the gap between their IL and the TL even though a TL model is provided to them" (p. 295).

Interestingly, although the advanced students generally verbalized more 'with-a-reason' LREs than 'noticing-only' LREs, the number of lexical 'with-a-reason' LREs produced by them was substantially smaller than their lexical 'noticing-only' LREs. This may indicate that even though an L2 learner notices certain aspects of language more frequently, that does not necessarily indicate that the learner notices

them with a higher level of awareness.

### 6.3. Data analysis for Research Question 3: *Is there any difference in the noticing between different types of writing tasks?*

The results for the third research question revealed that there were certain differences in quantity and quality of noticing between Task 1 (descriptive essay-writing) and Task 2 (argumentative essay-writing). As can be seen from the data, the model essay for Task 1 elicited more frequent LREs from the participants than the model essay for Task 2 despite the smaller number of words required in Task 1 (150 words vs. 250 words). One explanation is that the model for a descriptive essay is more resourceful as a feedback tool than the model for an argumentative essay. Another possibility is that the resourcefulness of Task 1 results from the difference in writing modes.

Some researchers have argued that the modes of writing can affect L2 writing performance (e.g., Dvorak, 1987; Koda, 1993). According to Schultz (1991), the types of writing are categorized into four modes of discourse – narration, description, exposition, and argumentation – and argumentative writing, in particular, has more cognitive demand and requires to be instructed. Reed, Burton and Kelly (1985) classify writing into three types – narration, description, and persuasion – and persuasion is the most difficult followed by description and narration, respectively. Furthermore, there has been some research which indicates that task difference can affect what learners notice in processing TL (e.g., Doughty, 1991).

The finding that Task 2's model essay provided the students with an opportunity to notice discourse more frequently might be related to the mode of writing. Task 2's

model essay enabled the students to notice issues of discourse, such as organization of paragraphs, inter-sentential relationship, and coherence and cohesion, with specific reasons. These results seem to indicate that the model essay for Task 2 may be beneficial not only for noticing on a surface level but also for noticing at a higher level of awareness. In other words, it seems that the mode of ‘argumentation’ can facilitate L2 writers’ ability to understand these discourse matters more deeply. Similarly, a possible explanation for the finding that Task 1’s model essay was useful for noticing form issues such as grammar, punctuation and spelling with a high level of awareness, may be that a careful examination of the writing mode of ‘description’ may be useful for understanding these formal aspects of language in L2 composition.

The validity of the discussion above can be confirmed with some of the interview data. In response to the second question of the interview (“Which model essay is the more beneficial one for you, Task 1 or Task 2? Why?”), nine students (six university students and three EAP students) reported that the model essay for Task 2 was more beneficial for them. They explained that Task 2’s model was useful for familiarizing themselves with argumentation, but that was not the case with Task 1. The remaining students (two university students and three EAP students) reported that Task 1’s model was more beneficial to them, as describing the table involves certain conventions such as special expression, comparative and superlative forms, and appropriate tense (i.e., form aspect) as well as how to select the required information.

## CHAPTER 7

### Conclusion

#### 7.1. Summary of the Study

The present study examined the role of noticing in comparing students' own texts with model essays in the IELTS writing test. The findings of the study revealed that the model essays led the L2 learners to notice various aspects of language. The findings also indicated that the learners' proficiency levels and the differences of the writing tasks were significant factors in explaining the difference in quality and quantity of their noticing. The analysis of the level of the learners' awareness revealed that the learners' frequent noticing of a certain aspect of language did not necessarily mean the noticing was at a high level of awareness. Moreover, the contrastive analysis of the current study and the previous relevant studies revealed that model essays not only functioned in a similar way to the texts reformulated by native English teachers, but also had unique advantages as feedback tools.

#### 7.2. Pedagogical Implications

Based on the findings, several implications for L2 writing pedagogy, especially EAP courses, can be drawn. First, there are certain language aspects which L2 English learners can notice only by using model essays and cannot notice with reformulated texts, such as paragraphs, punctuation, and content. In the reformulation approach of writing instruction, it is difficult for L2 writers to deliberate on the content of compositions. With this obvious advantage of model essays, it is important for English teachers to encourage students to think about what to write. The feedback for 'content' is crucial for L2 writers, as a major criterion of writing assessment is 'content'.

Second, it might be more beneficial for L2 writers to refer to various types of model essays. In response to the third question in the interview (“Besides model essays, what do you need in order to obtain more beneficial feedback? Please describe what, if anything, this model essay approach lacks.”), the students presented a wide variety of suggestions. Seven students (five university students and two EAP students) reported that they would like to refer to more model essays so as to know the examples of poor, average, and excellent examples as sample answers. Some EAP students reported that the model essay adopted in the current study was too advanced and showed interest in having a look at the sample essays which are at a slightly more advanced level than their own essays.

The third implication is the necessity of L2 writing instruction. As the results of the first research question revealed, some problems of language were difficult for certain students to solve by themselves without instruction. The results for the second and third research questions also suggested that the tendency of students to either notice or not notice a particular aspect of language depended on the learners’ proficiency level and the types of writing task. Likewise, whether learners notice at a superficial level or understand deeply at a high level of awareness also depended on L2 proficiency and task differences. Thus, instruction in a small study-group context could be essential especially for learners with the lower level of L2 proficiency, as five EAP students commented that it would be more helpful if an English teacher guided them in focusing on each aspect of language.

There were also several drawbacks involved in using a model essay as a feedback tool. Although some students reported that they would like to use model essays as exemplars against which they could judge the quality and content of their own

writing, that would not always be easy if the model essay were totally different from students' essay. In fact, in response to the first question of the interview ("Do you think the all activities were beneficial to improve your writing?"), 12 students answered "yes". However, two students answered negatively, reporting that the model essay was not beneficial as they could not find any similarities between their own essays and the model essay (especially Task 2's model essay).

For efficient and effective L2 writing instruction, the use of model essays should be encouraged in EAP courses. It is an efficient use of students' study time and will promote students' noticing as well as draw their attention to various aspects of L2 composition. It is also highly possible especially for more proficient L2 writers to take advantage of model texts in a self-study environment. In the classroom context, teachers should carefully select model essays suited to their students' level of L2 proficiency. Furthermore, *how* teachers could promote students' noticing a certain aspect of language and *how* they could improve the quality of noticing will be crucial issues in L2 writing pedagogy.

### 7.3. Limitations of the Study and Suggestions for Further Research

Given the small number of the sample size, it is difficult to extrapolate the results from the current study to other cases of L2 writing. In addition, because the two model essays adopted in the study may not be representative of all descriptive or argumentative essays, the generalization of the finding pertaining to task differences is also problematic. Ideally, this type of L2 writing study will be repeated with different model essays to confirm the results of the present study.

Furthermore, the argument that 'noticing only' can directly indicate a low level

of awareness and ‘accepting model with a reason’ equates to a high level of awareness requires further discussion, as Sachs and Polio (2007) suggest. According to them, Schmidt’s (2001) argument that noticing and understanding can be separated according to the level of awareness has been empirically supported in the relevant research. However, despite the consistency between the theoretical argument on noticing and the supporting evidences in the relevant empirical research (e.g., Qi & Lapkin, 2001; Hanaoka, 2007), there are still problematic issues in the operationalizing of ‘noticing’. For instance, the students might have had difficulties describing a specific reason in verbalizing LREs even though they were ‘understanding’ particular aspects of language due to the difference in the nature of each aspect of language. Ideally, in adequate consideration of these limitations, further research will conduct experiments in the pre- and post-test design using a larger sample, in order to investigate whether model essays effectively enhance L2 writing skills or not.

## References

- Astika, G. G. (1993). Analytical assessments of foreign students' writing. *RELC Journal*, 24(1), 61-72.
- Bowles, M. A., & Leow, R. P. (2005). Reactivity and type of verbal report in SLA research methodology: Expanding the scope of investigation. *Studies in Second Language Acquisition*, 27(3), 415-440.
- Cohen, A. (1987). Student processing of feedback on their compositions. In A. Wenden & J. Rubin (Eds.), *Learner Strategies in Language Learning* (pp. 57-69). Englewood Cliffs, NJ: Prentice-Hall.
- Coley, M. (1999). The English language entry requirements of Australian universities for students of non-English speaking backgrounds. *Higher Education Research and Development*, 18, 7-18.
- Collins, A., & Gentner, D. (1980). A framework for a cognitive theory of writing. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing: An interdisciplinary approach* (pp. 51-72). Hillsdale, NJ: Erlbaum.
- Collins, A., Neville, P., & Bielaczyc, K. (2000). The role of different media in designing learning environments. *International Journal of Artificial Intelligence in Education*, 11, 144-162.
- Cumming, A. (1995). Fostering writing expertise in ESL composition instruction: Modeling and evaluation. In D. Belcher & G. Braine (Eds.), *Academic writing in a second language: Essays on research & pedagogy* (pp. 375-397). Norwood, NJ: Ablex.
- de Bot, K. (1996). The psycholinguistics of the output hypothesis. *Language Learning*, 46(3), 529-555.

- Doughty, C. (1991). Second language instruction does make a difference: Evidence from an empirical study of relativization. *Studies in Second Language Acquisition, 13*, 431-469.
- Dvorak, T. (1987). Is written FL like oral FL? In B. Van, R. Dvorak & F. Lee (Eds.), *Foreign Language Learning: A research perspective* (pp. 79-91). Cambridge, MA: Newbury House.
- Ellis, R. (1993). Interpretation-based grammar teaching. *System, 21*, 69-78.
- Ellis, R. (2001). Investigating form-focused instruction. *Language Learning, 51*, 1-46.
- Eschholz, P. A. (1980). The prose models approach: Using products in the process. In T. R. Donovan & B. W. McClelland (Eds.), *Eight approaches to teaching composition* (pp. 21-35). Urbana, IL: National Council of Teachers of English.
- Ferris, D. (1995). Student reactions to teacher response in multiple-draft composition classrooms. *TESOL Quarterly, 29*(1), 33-53.
- Ferris, D., & Hedgcock, J. S. (1998). *Teaching ESL composition: Purpose, process, and practice*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Goby, V. P. (1997). Arguments against providing model answers in the writing skills classroom: The Singaporean case. *TESL Reporter, 30*(2), 28-33.
- Hanaoka, O. (2007). Output, noticing, and learning: An investigation into the role of spontaneous attention to form in a four-stage writing task. *Language Teaching Research, 11*(4), 459-479.
- Hedgcock, J. S., & Lefkowitz, N. (1994). Feedback on feedback: Assessment learner receptivity to teacher response in L2 composing. *Journal of Second Language Writing, 3*(2), 141-163.
- Hyland, K. (2003). *Second language writing*. Cambridge: Cambridge University Press.

- Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition*, 24(4), 541-577.
- Judy, S. (1980). The experiential approach: Inner worlds to outer worlds. In T. R. Donovan & B. W. McClelland (Eds.), *Eight approaches to teaching composition* (pp. 37-51). Urbana, IL: National Council of Teachers of English.
- Koda, K. (1993). Task-induced variability in FL composition: Language-specific perspectives. *Foreign Language Annals*, 26(3), 332-346.
- Leki, I. (1990). Coaching from the margins: Issues in written response. In B. Kroll (Ed.), *Second Language Writing*. New York: Cambridge University Press.
- Leow, R. P., & Morgan-Short, K. (2004). To think aloud or not to think aloud: The issue of reactivity in SLA research methodology. *Studies in Second Language Acquisition*, 26(1), 35-57.
- Levenston, E. A. (1978). Error analysis of free comparison: The theory and the practice. *Indian Journal of Applied Linguistics*, 4(1), 1-11.
- Murray, D. M. (1980). Writing as process: How writing finds its own meaning. In T. R. Donovan & B. W. McClelland (Eds.), *Eight approaches to teaching composition* (pp. 3-20). Urbana, IL: National Council of Teachers of English.
- Porte, G. (1996). When writing fails: How academic context and past learning experiences shape revision. *System*, 24(1), 107-116.
- Qi, D. S., & Lapkin, S. (2001). Exploring the role of noticing in a three-stage second language writing task. *Journal of Second Language Writing*, 10, 277-303.
- Reed, M., Burton, J., & Kelly, P. (1985). The effect of writing ability and mode of discourse on cognitive capacity engagement. *Research in the Teaching of*

*English, 19*, 283-295.

- Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning, 45*(2), 283-331.
- Robinson, P. (2003). Attention and memory during SLA. In C. Doughty & M. Long (Eds.), *Handbook of Second Language Acquisition* (pp. 631-678). Oxford: Blackwell.
- Russo, J. E., Johnson, E. J., & Stephens, D.L. (1989). The validity of verbal protocols. *Memory & Cognition, 17*, 759-769.
- Sachs, R., & Polio, C. (2007). Learners' uses of two types of written feedback on a L2 writing revision task. *Studies in Second Language Acquisition, 29*(1), 67-100.
- Scarcella, C., & Oxford, L. (1992). *The tapestry of language learning: The individual in the communicative classroom*. Boston, MA: Heinle and Heinle.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158.
- Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 1-63). Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). New York: Cambridge University Press.
- Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237- 326). Rowley, MA: Newbury House.

- Schultz, J. M. (1991). Writing mode in the articulation of language and literature classes: Theory and practice. *The Modern Language Journal*, 75(4), 411-417.
- Scovell, D., Pastellas, V., & Knobel, M. (2004). *404 Essential tests for IELTS: Practice tests for IELTS: Academic module book*. Sydney: Adams & Austen Press.
- Smagorinsky, P. (1989). The reliability and validity of protocol analysis. *Written Communication*, 6(4), 463-479.
- Smagorinsky, P. (1992). How reading model essays affects writers. In J. Irwin & M. Doyle (Eds.), *Reading/writing connections* (pp. 160-176). Newark, DE: International Reading Association.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.
- Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review*, 50(1), 158-164.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook, & B. Seidhofer (Eds.), *Principles and practice in applied linguistics* (pp. 125-144). Oxford: Oxford University Press.
- Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64-81). Cambridge: Cambridge University Press.
- Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471-483). Mahwah, NJ: Lawrence Erlbaum Associates.

- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16(3), 371-391.
- Swain, M., & Lapkin, S. (2002). Talking it through: Two French immersion learners' response to reformulation. *International Journal of Educational Research*, 37, 285-304.
- Thornbury, S. (1997). Reformulation and reconstruction: Tasks that promote "noticing". *ELT Journal*, 51, 326-335.
- Tickoo, A. (2001). Re-examining the developmental sequence hypothesis for past tense marking in ESL: Transfer effects and implications. *Prospect: A Journal of Australian TESOL*, 16(1), 17-35.
- Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, 16(2), 183-203.
- UCLES. (2007). *IELTS handbook 2007*. Retrieved from the IELTS Web site: [http://www.ielts.org/pdf/IELTS\\_Handbook\\_2007.pdf](http://www.ielts.org/pdf/IELTS_Handbook_2007.pdf)

## Appendix A

## WRITING TASK 1

You should spend about 20 minutes on this task.

Tourism in the United Kingdom contributes billions of pounds sterling to the UK economy. The table below identifies the twelve most visited paid-attractions in the UK in 1999 and 2000.

Write a report describing the information shown in the table.

You should write at least 150 words.

## The Top Most Visited Paid-Admission Attractions in the UK 1999 &amp; 2000

Attraction	London	Country	Visits 2000	Visits 1999	% Change
Millennium Dome	London	England	6,516,874E	Not opened yet	N/A
British Airways London Eye	London	England	3,300,000E	Not opened yet	N/A
Alton Tower	Alton	England	2,450,000	2,650,000	-7.5
Madame Tussaud's	London	England	2,388,000	2,640,000	-9.5
Tower of London	London	England	2,303,167	2,428,603	-5.2
Natural History Museum	London	England	1,577,044	1,696,725	-7.1
Chessington World of Adventures	Chessington	England	1,500,000E	1,550,000	-3.2
Legoland Windsor	Windsor	England	1,490,000	1,620,000	-8.0
Victoria & Albert Museum	London	England	1,344,113	1,251,396	7.4
Science Museum	London	England	1,337,432	1,483,234	-9.8
Flamingo Land Thema Park & Zoo	Kirby Misperton	England	1,301,000E	1,197,000E	8.7
Canterbury Cathedral	Canterbury	England	1,263,140E	1,318,065E	-4.2

E = Estimated

N/A = Not Applicable

## WRITING TASK 2

You should spend about 40 minutes on this task.

Present a written argument to an educated reader with no specialist knowledge of the following topic:

Discuss when, if ever, capital punishment can be viewed as a valid punishment for crime.

You should use your own ideas, knowledge and experience and support your arguments with examples and relevant evidence.

You should write at least 250 words.

## Appendix B

### Model Essay for Task 1 and Task 2

#### Task 1

The table in Task 1 identifies the twelve most popular tourist attractions in the United Kingdom in 1999 and 2000.

The table clearly shows that the two newest paid-admission attractions, The Millennium Dome and The British Airways London Eye, were the most popular in 2000 with an estimated 6,516,874 and 3,300,000 visitors respectively.

In 2000, all of the other attractions with the exception of The Victoria and Albert Museum and Flamingo Land had lower numbers than in 1999. The most significant decrease was experienced at the Science Museum (-9.8%) followed closely by Madame Tussaud's (-9.5%). However, Madame Tussaud's still remains one of the top four UK paid attractions with over two million visitors in 2000. The least significant change was experienced at Chessington World of Adventures which had 1,500,000 paid visitors in 2000 and 1,550,000 in 1999.

There does not appear to be any correlation to falling numbers in city or regional areas.

## Task 2

To many people, capital punishment is abhorrent, but it is possible to view it as a valid means of punishment. In my opinion, it can be justified as a penalty when three conditions are present. In this essay, I intend to discuss these three issues in more depth.

The rules of society are made and enforced for the purpose of protecting its members. In cases where the murder that has been committed is so frightful that society cannot allow the perpetrator back into its fold, then the death penalty may be the only answer. For example, crimes against humanity such as genocide; terrorist attacks which have resulted in many lives lost.

If the offender shows no remorse at all and is incapable of feeling any shame for his terrible actions, the only solution may be death. In these cases, it is clear that the offender cannot be rehabilitated.

Another important consideration is to ensure to the best of everyone's ability that the accused is guilty beyond reasonable doubt. If someone is punished with imprisonment, and later found to be innocent, he can be released from prison. However, when the punishment is death, it is impossible to make amends. There have been a number of examples in history where a man was found to be innocent after being hanged.

In conclusion, I believe that it is important to consider three main issues. The crime must be so bad that it would be impossible to welcome the perpetrator back into society. There must be no hope that the perpetrator can ever be rehabilitated back into society. There must be no doubt as to the perpetrator's guilt. In any situation that does not contain these components, I believe it is morally wrong to impose the death penalty.