Second Language Acquisition of Japanese Orthography

Yoko Matsumoto-Sturt

A thesis submitted in fulfilment of requirements for the degree of
Doctor of Philosophy

to
Theoretical and Applied Linguistics
The School of Philosophy, Psychology & Language Science,
University of Edinburgh

May 2004

Abstract

The aim of this thesis is to address issues on second language (L2) acquisition of Japanese orthography. In particular, I investigate L2 acquisition of reading and writing in Japanese with special attention paid to how linguistic and perceptual difficulties influence lexical processes involved in reading and writing by English-speaking learners of Japanese.

I first establish that L2 spelling problems are found in writing by keyboard. Evidence is presented from error patterns of L2 language used in a small self-constructed L2 error corpus that was gathered according to originally set external criteria. Next, I demonstrate that there are predictable error causes as well as general linguistic and perceptual problems among L2 writers. Real-time observational data illustrates how and when L2 writers make spelling mistakes with the keyboard, and an analysis of verbal protocol data reveals L2 writing strategies. Third, I turn to the domain of L2 perception and L1-specific listening strategies. I show the different patterns of learners' perceived rhythmic units in terms of L2 Moraic Awareness of Japanese words, and determine the extent of the use of L1-specific listening strategies. Fourth, I provide a general picture of Interlanguage (IL) lexical representations in reading and spelling. Evidence from novel and existing experimental work is presented which shows that L2 writer's linguistic problems are reflected in their written products. I present an account of a typical learner strategy of sublexical reading and writing. Finally, in the domain of visual kanji recognition, visual attention is addressed. An originally defined phenomenon of 'kanji illusion' leads to the interesting result that linguistic factors are not solely responsible for failures to notice kanji errors. This represents a new kind of explanation for L2 kanji reading difficulties, from a psycholinguistic perspective.

Declaration

I hereby declare that this thesis is of my own composition, and that it contains no material previously submitted for the award of any other degree. The work reported in this thesis has been executed by myself, except where due acknowledgement is made in the text.

Acknowledgements

While I was writing up this thesis, envisaging this part is what kept me going this far, because I understood the enormous number of people who deserve credit for what follows, many of whom probably do not even get a mention here. I am especially indebted to the University of Edinburgh for its generous support and encouragement for me to complete my thesis while working as a full-time lecturer at Asian Studies, where my colleagues and students past and present constantly gave me their warm support. I want to thank them for their good wishes.

My gratitude also goes to my supervisors, Drs Hugh Trappes-Lomax, Mitsuhiko (aka. Mits) Ota, and Professor John Joseph who have provided large amounts of feedback, information and other assistance with this thesis. To Hugh, my main supervisor at Theoretical and Applied Linguistics, I am grateful for guiding me through all those years from a very early stage of my thesis project till this end. My prime target reader has always been Hugh, whose critical comments during our discussions from the viewpoint of a non-native speaker of Japanese have shaped the style of this thesis. Last time I met him, Hugh told me that he has started learning Japanese, including how to write in hiragana. It was my privilege to appreciate his sincere desire to understand the very core of my thesis. It goes without saying that this thesis has greatly improved as a result of another supervisor Mits who has read draft versions of my thesis, and gave me incisive line-by-line suggestions and comments on the phonological and Japanese linguistic aspects of my thesis. I thank them for being patient with me, allowing me to pursue an interdisciplinary path, and for providing me with thought-provoking commentary and sage advice.

I would also like to thank my examiners, Mary Flaherty and Vivian Cook, as well as my viva convenor Jim Hurford, who all made my viva a stimulating and congenial experience.

and *Japanese* tea for me when there were ups and downs. He has also read entire drafts of the thesis in great detail, raised many important questions and clarified many others. Without him, this thesis wouldn't have existed.

Finally, there are my families. I want to thank Ann and Harry Sturt for their warm support and tremendous understanding of their habitually unsociable daughter-in-law for many years. I also thank my parents Fumie and the late-Tastuo Matsumoto for their love, and both financial and emotional support. Even when no one including myself was so sure about whether I could stick with this thesis until it got done, it was my mother who believed in me. This is where my strength came from. I also thank my sister Kyoko, and my brother Hiroyuki for their constant support and many years of sacrifices on my behalf. 長い間、ありがとうございました (I owe them everything).

Contents

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

Reading and writing in Japanese has been an active area of research in the field of Japanese pedagogy, second language acquisition (SLA) (Flaherty & Noguchi, 1998; Hatasa, 2002; Horiba, 1996; Jorden & Chaplin, 1976; Shimizu & Green, 2002; Sugiura et al., 1997; Tateoka, 2001; Tomita & Sanada, 1991; Usami, 2001), and Japanese Psycholinguistics (see, Kess and Miyamoto (1999) for their comprehensive review on the past work in this area). The past two decades have also seen a rapid growth in our understanding of the effects of first language (L1) orthographic systems on second language (L2) reading processes and strategies (Chikamatsu, 1996; Koda, 1987, 1990, 1995; Mori, 1998; Shimamura, 1987). Unlike English, which is alphabet-based, and employs a single system of orthography, the Japanese orthography system is a mixed system of kana syllabaries (hiragana and katakana) and logographic kanji (Chinese characters). Thus, typical printed Japanese is written in what is called as 'kanji-kana mixed' style where kanji (e.g. content words) and hiragana (e.g. grammatical particles and inflectional endings) are blended in one sentence. Within L2 Japanese pedagogy, it has been well established that learning to read and write Japanese orthography, especially kanji, is an outstandingly difficult task for learners from a non-kanji background (Kawaguchi, Kano, & Sakai, 1995; Hadamitzky & Spahn, 1997; Hatasa, 1989; Henshall & Takagaki, 1990; Ishida, 1989; Matsunaga, 1995; Takebe, 1989; Sakade, 1959; Tollini, 1994).

However, since the digital age reached Japan over two decades ago, keyboard-based new literacy tools such as Japanese word processors, electronic dictionaries and PCs have brought a major shift in Japanese writing practice. That is, the traditional notion of kanji writing as 'reproducing the shape of kanji stroke by stroke' is no longer necessary when we write with the keyboard. Knowing that Japanese children

must receive at least nine long years of handwriting practice in their compulsory education in order to acquire about 2,000 kanji, it would be too naive to believe that the physical swap of pen for keyboard is all that is involved in the shift to the computer writing environment in Japanese. Without a doubt there are literacy educators who can see the enormity of the impact of digital writing on the traditional notion of literacy; especially, a true understanding of such an impact in terms of the relationship between digital writing tools and developing literacy will surely make for a rich vein of research to be mined. In L2 educational settings, the study of pedagogical implications of this issue has just recently begun, thus investigating related issues is well timed.

However, although this thesis is concerned with L2 spelling mistakes[1] written by keyboard, the main aim of this thesis is not to address instructional issues of Japanese word processors within the scope of classroom research. Rather, I intend to address these issues more indirectly, by considering the *general* questions relating to L2 acquisition of Japanese orthography by English-speaking learners of Japanese, but concentrating particularly on issues that are relevant to electronic literacy. In particular, I aim to investigate L2 acquisition of reading and writing in Japanese, paying special attention to how linguistic and perceptual difficulties influence the lexical processes involved in reading and writing by English-speaking (EJ) learners of Japanese.

Many L1 developmental researchers have empirically demonstrated that the development of orthographic knowledge is underpinned by phonological knowledge (e.g. Ehri (1987), Treiman (1993)), and I agree with this point. Nevertheless, it must be noted that there is one crucial difference between acquisition of phonology and orthography. As the existence of illiterate people can demonstrate, one's ability to speak a language does not necessary imply the ability to acquire the written form of the language without conscious effort. For example, a considerable amount of exposure to print has been seen to be necessary to foster the acquisition of spelling development. Furthermore, spellings are frequently not phonetic, and the writers need to perform sound-letter mapping according to stipulated writing conventions. In other words, at least L1 phonology is largely acquired through unconscious learning, while the acquisition of orthography needs explicit learning and conscious knowledge. Given this, it is clear that the studies of phonology and orthography cannot always be treated as two sides of the same coin. Finally, it

---

[1]I use 'spelling' errors and 'orthography' errors equivalently throughout this study to refer to any incorrect writing involving individual kana letters or kanji characters.

is important to note that orthographic representations were developed for eyes. For example, many kanji words include sound cues and semantic information that are not available in the corresponding spoken forms, so that, for example, readers of kanji can visually differentiate the pair of homophone kanji compounds /ki.ki/, which also share an identical high-low (HL) pitch pattern, by their different visual forms (/ki.ki/ 危機 'crisis' and /ki.ki/ 機器 'equipment'), while the hearers need higher-level context to work out the meanings of such pairs. This means that the readers' visual kanji recognition skills play a crucial role in the reading comprehension of printed Japanese, because the Japanese language has an infamously large number of homophonic kanji words.

Chapter 3 takes up the problem of determining the course of this thesis. The study deals with Error Analysis of L2 spelling errors written using a Japanese word processor called JWP. I identify the pattern of learner errors, and show that the use of Japanese word processors somehow had a negative impact on L2 learners' spelling processes. Largely, there are two key areas involved in the potential causes of L2 writing difficulties. The first writing problem seems to be based on phonological difficulties for processing hiragana letters. The second type of spelling problem points to a perceptual problem of visual recognition for kanji processing. The result of this chapter leads to the conclusion that we cannot expect fruitful outcomes from research that deals with higher-level writing effects such as revision processes, without also considering the spelling problems of the individual words, because bottom-up and top-down processes are in constant interaction, and macro-level text interpretation is based on the knowledge of the individual building blocks of words. Thus, each study in the following chapters addresses a specific question relating to the general issue of why and how such L2-specific spelling errors occur, and why L2 writers often fail to notice their spelling errors. Accordingly, in chapter 2, I present background knowledge on the Japanese orthography system, along with relevant literature on Japanese pedagogy, computer-based writing, and speech and visual perception .

Chapter 4 serves to refine the factors which underlie the L2 spelling errors identified in chapter 3, and these in return generate research hypotheses for the following chapters. For this purpose, chapter 4 takes a qualitative approach; namely real-time observation of L2 writing processes and collecting verbal protocol data. These observational data are informative about real-time man-machine interaction and learner writing strategies, and illustrate how and when L2 writers make spelling

mistakes with the keyboard. In this chapter, I demonstrate that there are predictable error causes as well as potential linguistic and perceptual problems among L2 writers.

Chapter 5 turns to questions related to phonological awareness, which is generally considered to play a central role in literacy development. I look into L1-specific listening strategies with regard to the level of perceptual sensitivity to the internal structure (i.e. mora) of the Japanese syllable. A widely used L1 developmental methodology, the phonological tapping test, is adapted for this purpose. I show the different patterns of learners' perceived units, and identify the extent of the use of L1-specific listening strategies. Hence, this study presents perceptual data that helps us to interpret the following studies.

Chapter 6 addresses an overall issue of the quality of Interlanguage (IL) lexical representations in reading and spelling, for which I develop two reading tasks (i.e. the pronunciation task and the naming task), as well as a spelling task. In addition, the spelling task deals with an important SLA issue of L1 transfer, in terms of the strength of connection between L1 script-IL phonology vs. L2 script-IL phonology. In this study, I demonstrate that L2 writer's linguistic problems are reflected in their written products. The study shows how a particular linguistic or perceptual constraint can impose on L2 writing and reading difficulties, explaining the typical learner strategy of sub-lexical reading and writing. Three studies also illustrate a complete picture of acquisition order of a given set of kanji compound words by two different proficiency learner groups. The study also includes the first demonstration of frequency effects on naming latencies for L2 learners of Japanese.

Finally, in chapter 7, I address questions from the domain of visual kanji recognition and visual attention. The study examines the assumption that writers' kanji verification processes are affected by the degree of similarity between the error kanji and the target kanji, in terms of visual and contextual relatedness. I propose an original definition of the notion of 'kanji illusion', where the reader fails to detect a wrong kanji in the experiment, despite a subsequent demonstration that he or she has the correct lexical knowledge of the given kanji. This leads to the interesting result that linguistic factors are not solely responsible for the failure to notice kanji errors, thus providing an answer to the second general question of why L2 writers often fail to notice their kanji compounds errors. This represents a new kind of explanation for L2 kanji reading difficulties from a Psycholinguistic perspective.

CHAPTER 2

REVIEW OF THE LITERATURE

This study investigates second language (L2) acquisition of reading and writing in Japanese, with special attention paid to how linguistic and perceptual difficulties influence lexical processes involved in reading and writing by English-speaking (EJ) learners of Japanese. In this chapter, I provide background knowledge of the four types of Japanese orthography - two kana scripts, kanji characters and romaji as an essential starting point. I also summarize findings from three relevant literatures, in the fields of Japanese pedagogy, computer writing, and speech and visual perception as foundation to the following chapters.

2.1 Japanese orthography

Japanese orthography contains a combination of five different scripts: *kanji*, two versions of *kana*, Roman alphabet and Arabic numerals. Three primary scripts are kanji and two kana syllabaries ('plain kana' *hiragana* and 'partial kana' *katakana*), none of which is alphabetic. Kanji is the logographic component of Japanese orthography, which were borrowed from Chinese around 5th century AD (Tomita & Sanada, 1991; Taylor & Taylor, 1995). After adopting Chinese characters, Japanese gradually started using kanji only for phonemic purposes in order to express the pronunciation of native words, especially, proper names. In this way, kana or karina (Lit. 'temporary writing') was developed as the phonetic component of Japanese orthography around 9th century (Henderson, 1982; Shibatani, 1990; Taylor & Taylor, 1995). The difference between two versions of kana is that in formally, the square shape of kanji is retained in katakana more than in cursive hiragana. In Modern Japanese, katakana is mainly used to transliterate loan words other than Chinese, as well as being used to express onomatopoeia, while hiragana's main function is to

write grammatical items such as post-positional particles and inflectional endings known as *okurigana*. Hiragana has one more important role to indicate the pronunciation of a kanji by appearing on the right or above the given kanji according to the way the text is written; vertically or horizontally. This function of hiragana is known traditionally as *furigana* or technically as *rubi* (i.e. a *ruby* in English). Example 2.1 shows a sentence that contains all five scripts and furigana. The sentence is read as 'ani-wa juu nen mae ni ejinbara daigaku-de AI-wo benkyoo-shita' (My elder brother studied AI at University of Edinburgh ten years ago):

Table 2.1: An example of Japanese sentence

Furigana	あに			ねんまえ			だいがく			けんきゅう	
	兄	は	10	年前	に	エジンバラ	大学	で	AI	を	研究	した。
Kanji (K)	K			KK			KK				KK	
Hiragana (h)		h			h			h		h		h h
Katakana (k)						k k k k k k						
Roman alph. (a)									aa			
Arabic num. (n)			nn									

(Note: Example sentence showing positions of kanji (K), hiragana (h), katakana (k), Roman characters (a), and Arabic numerals (n))

2.1.1 *Hiragana and speech sounds*

Hiragana scripts

The latest kana convention, Gendai kanazukai (Contemporary kana usage) which was established in 1946, is said to have a closer sound-letter correspondence in accordance with contemporary spoken Japanese (Ogawa, Hayashi, et al., 1982; Morita, 1989; Shibatani, 1990). The preface of Gendai kanazukai (henceforth 'the new hiragana rules') states that its aim is to give guidelines for laws and ordinances, official documents, newspapers, magazines, and broadcasting, and so on. However, the guidelines are not prescribed rules for what individuals or specialists in the field ought to do (Bunkacho, 1991). It is also clearly states that dialects, foreign loan words, and onomatopoeic expressions are excluded, because we transcribe foreign loan words and onomatopoeia with katakana letters under slightly different

katakana conventions. The new hiragana rules have an appendix that is a comparative chart of the new hiragana rules and Rekishi-teki kanazukai 'Historical kana usage' (henceforth 'the old hiragana rules').[1]

Hiragana and speech sounds

A mora is a single unit of timing (See Ladefoged (1982)). 46 basic kana letters and 2 types of diacritics represent 106 morae [2] in the abstract standard Japanese based on the Tokyo dialect. Thus, learning 46 kana letters means knowing all the moraic sound combinations of Japanese, which are 106 altogether and which are expressed in the traditional kana table (see appendix A) with 104 kana letters.

The basic consonant-vowel (CV) based kana letters contain a set of 5 vowels, including /a, i, u, e o/.[3] These produce 35 CV-type syllables in combination with 4 voiceless obstruents /k, s, t, h/, 2 nasals /n, m/, and 1 approximant /r/. There is also a syllabic nasal /N/. In addition, there are two semi-vowels /j, w/, which produce a further 4 semivowel-vowel combinations. An additional 25 kana letters are represented by two types of diacritic marks: *dakuten* or voicing dots (e.g. ベ), and *maru* or small circle (e.g. ペ). The voicing dots create 20 morae beginning with /g, z, d, b/, and the small circle produces 5 morae beginning with /p/. Three reduced size semivowel symbols や[ja],[4] ゆ[ju] よ[jo] make 12 voiceless and 9 voiced palatalized morae when they are combined with corresponding hiragana letters き (ki), し (si), ち (ti), に (ni), ひ (hi), み (mi), and り (ri). For instance, きゃ[kja]. The use of the two diacritics creates 12 further voiced counterparts (e.g. ぎゃ) and 3 special voiceless versions of consonant-semivowel type morae (e.g. ぴゃ).

Each hiragana letter has a one-to-one grapheme-mora correspondence, whereas there is often no such straightforward one-to-one correspondence between letters and sounds when a word is transcribed by a combination of hiragana letters. For example, the accusative particle を [o] is orthographically distinguished from お [o] in content words, while hiragana letters は ([ha] or [wa]) and へ ([he] or [e]) remain as a homograph of the topic particle *'wa'* and a directional particle *'e'*. Further

---

[1]Rekishi-teki kanazukai (the old hiragana rules) were not concerned with the discrepancy between contemporary pronunciations of the Edo period (1600-1868) and orthographic rules based on mid Heian period (794-1185). They were adopted by the government during the Meiji period (1868-1912), and survived to the end of World War II.

[2]Otake, Hatano, Cutler, and Mehler (1993) identify the number of phonologically permissible morae as 108. This number does not correspond to the number of morae in kana charts.

[3]I use the slash brackets (/ /) to denote phonemic transcription.

[4]I use square brackets ([ ]) to denote phonetic transcription.

more, when the particle *'wa'* appears at the end of a sentence with a different function beside as the topic marker, this particle is always written with わ [wa]. Such one-to-two grapheme-mora correspondences were created for historical reasons.[5] As for L2 learners, they would never properly communicate in written Japanese if they did not know these orthographic conventions. In practice, L2 beginners learn these writing conventions with some concrete examples as soon as they start learning hiragana.

Next, the usage of two pairs [(d)ʑi] and [(d)zɯ].[6] The pair じ (usually written in Romaji as "zi") and ぢ (usually written "di" in Romaji) share the same phonetic value [ʑi], and another pair ず (usually written "zu") and づ (usually written "du") are homophones of [zɯ]. According to Shibatani, historical and dialectal evidence show that the above pairs have had distinct phonetic values in the past (Shibatani, 1990, 165). Indeed, the old hiragana rules clearly provide four voiceless and voiced pairs: し [ɕi]- じ [ʑi], ち [tɕi]- ぢ [ʑi], す [sɯ]- ず [zɯ] and つ [tsɯ]- づ [zɯ], while the new hiragana rules reduces them into two pairs: i.e. し [ɕi]-じ (ぢ) [ʑi] and す [sɯ]-ず (づ)[zɯ] by deciding that ぢ (di) and づ (du) in the old version should be written as じ /zi/ and ず/zu/ respectively in the new hiragana rules. For instance, 'taste' - 味 [adʑi] is transliterated as あぢ with the old rules and あじ with the new rules. Although the new hiragana rules look simpler than the old ones, a closer look at the hiragana chart (see appendix) soon reveals that Japanese orthography still contains ぢ (di) and づ (du) as the voiced versions of ち/ti/ and つ/tu/ respectively.

The new hiragana rules explain that in order to improve the earlier version, ぢ (di) and づ (du) are preserved for two exceptions (Bunkacho, 1991, 202-205). Firstly, in the case of compound kanji, the second element of the compound goes through a sound change (sequential voicing). In other words, ぢ (di) or づ (du) should be used when the initial consonant of the second element in kanji compounds are morphologically related to the voiceless counterparts ち/ti/ or つ/tu/. For example, the underlined sounds of 'nosebleed' はな ぢ[hanadʑi] and 'crescent moon' みか づ き [mikadzɯki] are /hana.ti/ and /mika.tuki/ including 'blood' 血 (ち) /ti/ and 'moon' 月 (つ き) /tuki/ respectively, thus はな じ and みか ず き are not acceptable. Secondly, if a morpheme contains a sound sequence of [tsɯ-zɯ], or [tɕi-ʑi], such sound sequence should be written as つづ or ちぢ respectively. For instance, 'shrinkage' 縮み [tɕidʑimi] and 'Japanese hand drum' 鼓 [tsɯdzɯmi] should be written as ちぢ

---

[5]There have been three major kana orthography conventions in Japanese since the medieval period *Teika kanazukai*, and the current hiragana conventions follow the old hiragana rules with a few minor changes.

[6][ʑi] when appearing word-internal position, while [dʑi] at word-initial position.

み/tidimi/ and つづみ/tudumi/.

2.1.2 L2 acquisition of Japanese orthography

L2 Hiragana teaching and scope of the problem

Standard tables exist for kana, in which all the kana characters are organized in terms of 5 types of morae: V, CV, CCV, nasal coda, and geminate consonant. Kana is easy to learn for both first and second language learners of Japanese, because there is almost a one-to-one correspondence between symbols and sounds, and the shape of each letter is simple having just one to six strokes. Japanese children master kana characters by the end of the first grade (Sasanuma, 1975), but it is not unusual to see a pre-school child who can read and write hiragana. At Edinburgh university, first year students with no previous knowledge of Japanese are expected to learn both hiragana and katakana within a month of their study. However, it is well known in L2 orthography teaching of Japanese that some inconsistencies between orthography and expressed sounds bring about a variety of learners' writing errors. (Morita, 1989; Tomita & Sanada, 1991; Kimura, Sakata, Kubota, & Kawamoto, 1989). For instance, a teachers' handbook suggests that じ/zi/ and ず/zu/ should be taught first at a beginners level as fundamental rules to express phonetic [(d)ʑi] and [(d)ʑɯ] without mentioning exceptional use of ぢ (di) and づ (du), which should be introduced at later stage with concrete examples of sequential voicing, so that the teacher could prevent L2 learners from confusing those homophones (Tomita & Sanada, 1991). But it is obvious that these exceptional cases (i.e. sound-letter inconsistencies) create a considerable amount of confusion for L2 learners of Japanese trying to transcribe such sounds in hiragana. Therefore, finding an effective and efficient way of learning and teaching orthographic conventions of hiragana is the key for a successful L2 orthography teaching in Japanese. It is worthwhile examining the subject more closely.

To begin with, the new hiragana rules give guidelines for the treatment of vowel length, since long vowels contrast with short vowels in Japanese (Ladefoged, 1982, 225). For instance, the durational contrast of short vowel [a] and long vowel [aː] can be seen in a minimal pair of 'aunt' [obasaɴ] and 'grandmother' [obaːsaɴ], which are written as おばさん o-ba-sa-ɴ and おばあさん o-ba-a-sa-ɴ respectively. The basic rule is that a long vowel is realized by repeating the same vowel as the preceding item in the case of [aː], [iː], and [uː], while second element of the long vowels [eː] and [oː]

have a one-to-two sound-letter correspondence. For this reason, the long vowels [eː] spelled either as *e-i* or *e-e*, and [oː] spelled either as *o-o* or *o-u*, are more problematic than other long vowels in L2 orthography teaching. However, as Kimura et al. (1989) points out, we can find very few examples of the *e-e* version, as in 'yes' え え [eː], while there are many examples of the *e-i* variant in the pronunciation of kanji compounds. For example, 'cinema' 映画 (えいが) e-i-ga [eːga]. This makes it easier for L2 learners to transcribe kanji words that include this particular long vowel.

The actual problem, then, seems to be with the long vowel [oː] as the new hiragana rules exclude words, which are etymologically related to historical ほ [ho] and を [o], from the rule that the second element of the long vowel [oː] is expressed with hiragana letter う (u). To be more precise, in the old hiragana rules, the second element of this particular long vowel was realized with a hiragana letter of ほ [ho] or を [o]. For instance, [koːri] 'ice' was transcribed as こほり 'ko-ho-ri'. Similarly, 'to-o-ka' ('ten days') was transcribed as とをか. Because of these historical reasons, the guidelines in the new hiragana rules recommend ほ [ho] in certain contexts, and を [o] in all contexts except when it is a grammatical particle を [o], should be replaced with お [o]. Accordingly, we now write these words as こおり 'ko-o-ri' and とおか 'to-o-ka' respectively, under the current hiragana conventions. As a result, the current hiragana convention requires users to use two different spellings for the long vowel [oː], depending on the word that is being written (for example *o-u* in おうむ [oːmɯ] 'parrot', and *o-o* in とおり [toːri] 'street'). Because phonetic [oː] does not include this etymological information, L2 learners of Japanese never know why some words including [oː] cannot be written with the default う (u)-version like many other words. L2 learners simply need to memorize these exceptional words, but they should, at least, be given a brief explanation of why this inconsistency has occurred. I have observed many writing errors due to the sound-letter inconsistencies described above. The present study will be very much concerned with this issue, and we will discuss it further in the later chapters.

Learning to read and write kanji

It is worth mentioning that kanji writing is a difficult element of Japanese orthography acquisition even for L1 learners of Japanese. For example, in an L1 study, Hatta, Kawakami, and Hatasa (1997) reported that Japanese college graduates still made many kanji errors in their informal writing, even after spending a considerable amount of time learning kanji in their formal education from age seven. In the L1

educational setting, a total of 1006 educational kanji are taught during the six years of elementary education from age seven to twelve. L1 learners are subsequently expected to learn a further 950 kanji during the latter part of compulsory education from age thirteen to fifteen. On the contrary, L2 kanji learners are generally taught around 200 to 500 kanji within a year of study at much faster rate of learning.

Both L1 and L2 kanji learners need to spend a lot of time and effort in memorizing the shape and pronunciation of kanji, but L1 learners are allowed to learn kanji over a longer period compared with that of L2 learners. In addition, L1 children seem to have at least two years of time between developing the sound-meaning knowledge of words and learning the shape and pronunciation of kanji at primary school. On the contrary, L2 kanji learners are normally required to remember three aspects of kanji at one time: (a) shape, (b) meaning, and (c) word pronunciation. With a question of whether this learning method is effective, some L2 scholars and educators have suggested that in the L2 situation the spoken language should be taught prior to the introduction of the written language (DeFrancis, 1977; Jorden, 1986; Matsunaga, 1995). For instance, Matsunaga (1995) suggested that L2 learners of Japanese whose mother tongue is an alphabetic language, would benefit from a time lag, though she cannot specify the duration of time, between introduction of oral/aural skills and that of reading/writing skills. I shall discuss this implication further in 2.3.3.

2.1.3   Three elements of kanji

Visual representation of kanji

In this section, we shall concentrate on the issue of visual complexity of kanji. Any comprehensive kanji dictionary has a huge kanji inventory of up to 50,000 kanji characters. However, Kaiho and Nomura (1983, 121) points out that it is misleading to draw comparisons with the small number of letters in alphabetic systems, as the relevant comparison should be between the number of kanji and the number of words. It is the number of components or elements of kanji that makes it feel that kanji are complicated. Tamaoka and Takahashi (1999)'s study provides direct evidence for this claim. They reported that the average of initiation time (2150 ms) to write a frequent and visually less complicated two-morpheme kanji compound was about three times more than that of infrequent irregular English words (715 ms). They attribute this delay to the arbitrary relationship between

orthography and phonology in kanji, and visual complexity (i.e. number of strokes) of kanji.

As we can see from any kanji dictionary, orthographic representations of kanji are not built in a random manner. In other words, many kanji share a common element, and they are sorted according to shared graphic configurations in the kanji dictionary. For example, a typical kanji dictionary provides a list of 214 basic kanji radicals, which are used as a component of a complex single kanji, with their stroke numbers. In a relevant kanji study, Kaiho and Inukai (1982) investigated how L1 Japanese readers perceptually group visual representations of kanji. They surveyed subjective judgments of Japanese university students about the visual features of the 881 educational kanji. Kaiho and Inukai (1982) concluded that the following seven kanji categories typically represent the L1 perceptual or psychological building blocks of kanji: (1) complexity (simple or complex), (2) regularity (arbitrary or regular), (3) unity (scatter or compact), (4) elongation (flat or long), (5) openness (close or open), (6) straightness (slant or straight), and (7) stability (unstable or stable).

A kanji can be simple or extremely complex in terms of its number of strokes. For instance, the simplest kanji 一 has only one stroke, whereas the most complex kanji is composed of four copies of the 16-stroke kanji 龍, and as a result, this kanji requires 64 strokes all together (Takebe, 1989, 13). From this example, it is obvious that simple kanji are easy to recognize, but the opposite may also be true that a complex visual figure sticks out itself among others. A question arise: how complicated should a kanji be to acquire this attribute? Kaiho (1979) gives the basic line as 13 strokes. According to him, kanji processing difficulty increases up to 13 strokes, and then the recognition process will become easier due to the visual complexity of kanji. This suggests that visual complexity of kanji can be used as a part of top-down discriminative strategy for kanji reading (Kaiho & Nomura, 1983).

In this respect, there is an interesting report in Chen and Yuen (1991) on the different visual perception of kanji between three groups of L1 Chinese children of Taiwan, Hong Kong, and the People's Republic of China (PRC). All of them learned kanji script, but the PRC children acquired the simplified versions rather than the traditional Chinese. Children were asked to pronounce given kanji words (pseudohomophones or non-words that share identical phonological information) and to judge whether or not they were real words. Taiwanese and Hong Kong children equally recognized real words and rejected nonwords, and the PRC children performed equally well in recognizing real words. However, the PRC children tended

to be more affected than the other groups by orthographically legal nonwords, and thus, they were less accurately rejecting these words as real words. This means that Taiwanese and Hong Kong children are more sensitive to the deviation of visual forms, while the PRC children have more tolerance for such deviations as a result of their experience of using simplified version of kanji that were created by reducing some parts of kanji elements from the original shape. This finding may suggest that visually simplified kanji, which always have a smaller number of strokes than the traditional kanji, are not necessarily helpful for visual discrimination purposes.

As we can see considerable evidence from L1 and L2 written errors, the visual complexity of kanji is likely to cause spelling difficulties in reproducing graphic shape. For instance, a typical L2 kanji hand-written error has an imperfect kanji shape, which either misses some elements of kanji or contains redundant elements such as a small dot. In the following example, I shall demonstrate how a dot-like element (i.e. 丶) can differentiate one kanji from another.

(1) 九 (*nine*)   vs.   丸 (*circle*)
(2) 水 (*water*)   vs.   氷 (*ice*)
(3) 犬 (*dog*)   vs.   太 (*fat*)

It is clear that the appearance of a small dot-like element discriminates a pair of visually similar kanji in (1), (2), and (3). Equally, only one missing line in 烏 (*crow*) will tell the reader that this kanji is different from 鳥 (*bird*). From the writer's point of view, this line is also vital to convey the right meaning. In sum, visual complexity of kanji need not always lead to processing difficulty; rather this could be used as a visual cue. Nevertheless, hand-written production needs to be accurate with respect to the number of strokes, so as not to convey a different meaning with an orthographically similar but semantically distant kanji.

Semantic representation of kanji

A classic categorization of kanji is called *rikusho* that features four categories of kanji (Takebe, 1991). These categories are pictograph (*shōkei moji*), ideograph (*shiji moji*), compound-semantic (*kaii moji*), and semantic-phonetic (*keisei moji*). Pictographic kanji were derived from the drawings of concrete objects, for example 山 to express ⌂⌂ (mountain). Three thousand years ago, all kanji were pictographs, so that the association between kanji characters and their meanings was transparent. Some abstract notions such as numerals and spatial relations are expressed

by ideographic kanji. Compound-semantic kanji typically contain more than one pictographic or ideographic kanji. For example, compound-semantic kanji 林 *woods* and 森 *forest* were created by doubling or tripling a single pictograph 木 *tree* respectively. Probably most complicated of all are semantic-phonetic kanji. Typically, these kanji consists of one element to denote a broad semantic category and one element for pronunciation. For instance, the combination of a semantic radical *gold* 金 and a phonetic constituent 同 [doː] resulted in a semantic-phonetic kanji *copper* 銅 [doː]. However, statistically, kanji in first three categories are small in number, whereas about 80% of Japanese kanji fall into the semantic-phonetic type (Kess & Miyamoto, 1999).

The role of a semantic radical in semantic-phonetic kanji   In kanji dictionaries, kanji characters are listed according to 214 basic kanji elements known as radicals. The simplest view is that the average L1 Japanese reader uses the semantic radicals as a meaning cue. This view of the role of radical is not incorrect, but it is entirely mistaken to think that the semantic radicals tell the readers the exact meaning of a given kanji. For example, while 木 (*tree*) indicates 'plant' in the upper part of 杏 (*apricot*), the left part of 杉 (*Japanese cedar*), and the lower part of 栗 (*chestnut*), their visual forms indicate only a partial meaning as being a kind of plant, so that the full lexical meaning of each kanji must be remembered individually. This means that the function of the semantic radicals would be to guide the readers to an approximate semantic or lexical field such as a nature group, an animal group, a human-emotion group, a tool group, a food group and so on.

The effect of a phonetic radical in semantic-phonetic kanji   Generally speaking, the phonetic element of a semantic-phonetic kanji is read in on-yomi or Chinese pronunciation. For example, three kanji 白 (*white*), 拍 (*beat*), and 泊 (*stay*) share the same on-reading /haku/. However, it is important to note that kanji pronunciations in on-yomi underwent naturalization processes, hence lost each distinction marked by the tones in Chinese, which lead to the existence of too many homophonic on-readings of kanji in Japanese. As a result, the Japanese kanji system lost its power to differentiate meanings of kanji by on-reading (Chinese reading). In other words, the phonetic component of a semantic-phonetic kanji as a phonetic cue cannot be a very informative source in spoken Japanese. From the writer's point of view, the right visual representation of each homophonic kanji must be retrieved via meanings. For instance, homophonic kanji /sei/ 西, 生, 世, and 正 must be written according to their meanings *west*, *creation*, *the world*, and *straight* respectively. The same is also applied to a set of semantic-phonetic kanji /sjoo/:

消, 沼, 鐘, 礁, and 星 must be retrieved by their meaning *disappearance, marsh, bell, reef,* and *star* respectively.

Semantic ambiguity and kanji pronunciation   A well-known visual effect of kanji is to resolve semantic ambiguity in homophonic kanji such as those shown above. A pair of kanji 橋 (*bridge*) and 箸 (*chopsticks*) share the pronunciation /hasi/. The meaning is directly obvious from their different visual forms. For the spoken production[7], the prosodic information (i.e. the change of syllable pitch) is used to specify the intended meaning: 'bridge' with a low-high (LH) pitch pattern, and 'chopsticks' with a high-low (HL) pitch pattern. Note, however, that some homophonic kanji also have an identical pitch pattern. For instance, a LH pitch pattern is used in a homophonic kanji pair /kekkoo/ 決行 (*decisive action*) and 欠航 (*cancellation of a ship*). This would create a real semantic problem, because the word pronunciation /kekkoo/ suggests two or more meanings with different kanji. Viewed in this light, we can conclude that the lexical prosody provides inconsistent aid to the hearer in spoken perception, thus, we often rely on additional top-down information from the discourse level to pin down the intended meaning of homophony.

Next, we shall consider a case of lexically distinct homograph /itiniti/ 一日 (*a day*) and /tuitati/ 一日 (*the first day of a month*). It is impossible to identify the pronunciation of the given kanji in isolation without context or appearance of a string of furigana (i.e. kanji pronunciation indicator, written in hiragana) as a pronunciation aid. We can find similar homograph samples in English such as wind (*blowing*) and wind (*a clock*). For such words, the reader needs context-level information to determine the right meaning or the right pronunciation. In other words, the semantic information and/or top-down higher-level information play a role in the phonological disambiguation process. L2 educators usually regard visually simple kanji like 一日 as easy learning targets, but we should note that even such kanji would cause processing difficulty for the L2 learners of Japanese by demanding extra semantic or phonological processing effort.

Sound representation of kanji

As discussed above, some kanji have a phonemic component as a sound cue, but kanji do not have a systematic sound representation, unlike kana. In comparison

---

[7]The examples are spoken words in Tokyo Japanese.

with one reading for one kanji in Chinese, a single kanji character in Japanese typically has multiple readings of Chinese reading *on-yomi* and Japanese reading *kun-yomi*[8]. While it is not always the case, a single kanji is often read in kun-yomi and kanji compounds are normally read in on-yomi. As for kanji compounds, there are four reading combinations: (a) on-on, (b) kun-kun, (c) on-kun[9], and (d) kun-on[10] (Ogawa et al., 1982; Ito, 1989; Ishida, 1989). Although kanji compounds are most likely to be read in on-on combination, there is no systematic rule for determining the right pronunciation for a given kanji (Haratsuchi, 1989; Takebe, 1989; Taylor & Taylor, 1995).

There are two more readings (Taylor & Taylor, 1995, p.302): i.e. *ateji* and *jukuji-kun*, which may cause further reading complications to the sound representation of kanji. An ateji ignores meaning totally and uses the sound that is assigned to Japanese native word or loan words (Ogawa et al., 1982; Ito, 1989; Matsunaga, 1995). A typical example of ateji is 'Canada' 加 [ka] 奈 [na] 陀 [da]. However, in L2 kanji acquisition, the real problem lies with learning jukuji-kun (Idiomatic kun) that are made by adapting only the meaning of the given kanji and replacing their usual reading with a *wago* (native word) pronunciation, thus, they have a unique kun-reading. This means that a reading strategy, which employs sound clues from the constituent kanji in jukuji-kun compounds, always ends up with the wrong pronunciation. For instance, 'adult' *otona* (大人) is a combination of 'big' *dai* and 'person' *jin*. L2 readers often read this compound as *daijin* as a result of breaking this compound into the sub-lexical values of each element kanji in search of the most likely pronunciation of 大人. In *Jyōyō* kanji list, 110 commonly used jukuji-kun are listed in an appendix, from which I have identified 22 jukuji-kun[11] that are taught in the beginners class at Edinburgh.

2.1.4   Compound kanji

A single kanji can represent a word or it can appear as an element of a two- or more kanji compound. Here, we define a *kanji compound* as a lexical item that is composed of two or more kanji, and a *kanji word* as a lexical item that is composed of one or more kanji. However, the vast majority of lexical items in Japanese

---

[8]Of the 1,945 Jōyō kanji list, 737 (37.9%) kanji have only on-yomi, 40 (2.1%) kanji have only kun-yomi, and 1,168 (60%) kanji have both on and kun yomi. See Ito (1989, p.131).

[9]Traditionally known as *jūbako yomi* (重箱読み).

[10]Traditionally known as *yutō yomi* (湯桶読み).

[11]These are: 明日, お母さん, お父さん, 大人, 昨日, 今日, 今朝, 今年, 上手, 七夕, 一日, 時計, 兄さん, 姉さん, 二十, 二十日, 一人, 二人, 二日, 下手, 部屋, and 八百屋.

are compound kanji; to be more precise, they are said to represent 50% to 70% of the inventory of Japanese dictionaries (Tamaoka & Takahashi, 1999; Kess & Miyamoto, 1999). A great deal of research has been done on the issue of how these kanji are organized in the Japanese mental lexicon. Are they organized at the word level (whole-word view) or are they structured according to each constituent character (sub-lexical view)? First of all, researchers focused on the role of the first element of a compound kanji (Saito, 1981; Kaiho & Nomura, 1983; Hirose, 1992; Tamaoka & Takahashi, 1999). For instance, Hirose (1992) used a priming technique to investigate how the first kanji is used in the retrieval process. He concluded that the role of the first kanji seems to be that it is used in the initial stage of compound kanji identification as a search cue. He further suggested that the compound kanji are sorted in our mental lexicon according to the meaning of its first element. Saito (1981), Kaiho and Nomura (1983), and Tamaoka and Takahashi (1999) also came up with a conclusion similar to Hirose (1992).

The important point to note is that none of their conclusions support the decomposition of two constituents in a kanji compound, nor their separate identification. Rather, most of the literature on this issue appears to provide evidence in favor of the whole-word view, namely that kanji compounds are identified at whole word level rather than on the level of the individual character (Kaiho & Nomura, 1983; Hirose, 1992; Morton, Sasanuma, Patterson, & Sakuma, 1992; Nagae, 1992; Wydell, Butterworth, & Patterson, 1995; Sakuma, Sasanuma, Tatsumi, & Masaki, 1998; Tamaoka & Takahashi, 1999). The whole-word framework predicts limitations for learning kanji only by isolation forms and pronunciations, because there is no way for L2 learners to know the sound combination of a given kanji compound. In addition, the lexical meaning of a kanji compound cannot be assumed from the each constituent kanji of the kanji compound. These two points strongly suggest an effective L2 kanji learning should be seen as part of lexical acquisition.

2.2 Japanese word processing

2.2.1 Text input methods

Input Method Editor (IME)

In order to read and write Japanese text on computer, we need either a Japanese language operating system (e.g. Japanese Windows) or a Japanese language extension (e.g.TwinBridge Japanese Partner) that can handle Japanese input within an

English language operating system. In addition, there is a downloadable freeware Japanese word processor called JWP produced by Stephen Chung. Commercially available input method editors [12] (IME or IM) such as MS-IME, VJE, WXG, and ATOK allow us to convert keystrokes to Japanese characters. Although the specifications and efficiency are different, every application shares the general principles - 'kana (or romaji)-to-kanji conversion method' that does a dictionary lookup to convert hiragana into kanji. First, a writer sees a sequence of hiragana appearing on the screen (i.e. pre-edit window) while typing in hiragana letters according to the pronunciation of Japanese words. Next, the writer presses a convert button, and a kana chunk called *bunsetsu* will be converted to a kanji-kana mixed text after the built-in kana-to-kanji conversion dictionary has been searched for the target kanji or kanji compound. The following flow chart explains the conversion mechanism:

kana input → target chunk being selected → syntactic analysis

kana-to-kanji conversion ← database being searched ←
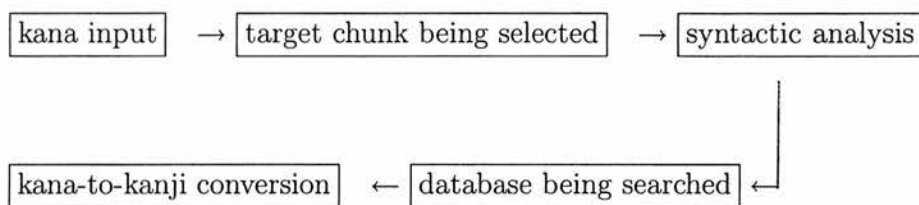
(Cited from Mori and Yagihashi (1989, 68) and translated by Matsumoto-Sturt)

Here, it is important to note that Roman characters and kana letters are always laid out within the same keyboard for the user to handle bilingual input. Thus, a writer has a choice between romaji nyūryoku (Roman alphabet input method) and kana nyūryoku (kana input method) even for writing Japanese sentences. The advantage of romaji nyūryoku is the use of the so-called QWERTY keyboard, which was designed for the English typewriter. This input mode requires users to remember many fewer key positions than the kana nyūryoku. On the other hand, mora-based kana input mode needs less frequent typing. For instance, when the kanji for 'cherry tree' 桜 [sakura] is required, a string of hiragana letters さくら appears in the pre-edit window in the document as the writer types in three hiragana letters さ (sa)-く (ku)-ら (ra) with the kana input mode, or six Roman characters 's-a-k-u-r-a' with the romaji input mode. Then for both modes, the writer needs to choose the desired kanji, hitting the return key to inform the IME that the writer recognized the correct kanji. According to Umemura et al. (1995, 172), L1 adult users prefer romaji nyūryoku because of the easy acquisition of keyboard operation skills.

---

[12]IMEs exist for other Asian languages such as Chinese, Korean, and Vietnamese.

A typical built-in dictionary within IME can remember the last choice as well as a user's tendency to use a specific area of the lexicon. This helps to minimize the processing speed of homophones when they are reused. For instance, if a user is a student of Economics, it is natural to assume that 'investment' 投資 [toː ɕi] should be more frequently used compared with an alternative homophone 'see-through' 透視 [toː ɕi]. This idea is known as a 'frequency learning system', in which most frequently used word appears as the first item in the kanji list according to the automatic frequency analysis of the used words. In addition, we can normally register any kanji / kanji compounds, which are not available from the built-in dictionary, in a 'user dictionary'. A better-performed IMEs feature a more sophisticated built-in phrase dictionary that can carry out a further analysis such as part of speech analysis and syntactic analysis, so that correct collocations are more likely to come out (Nikkei PC21, 1998). For example, さ (sa) く (ku) could be converted into homophones /saku/ 咲く ('bloom'), 裂く ('tear'), and 割く ('break'). If the writer is working on a sentence including the name of a flower, then a more advanced built-in dictionary can work out which homophonic verbs likely to take a flower as the subject. In this way, recent versions of IME offer more accurate and less time-consuming kana-to-kanji conversion, and users can expect near user-free homophone conversions at a more complicated sentence level.

Limits of J-word processing packages and the writer's role

From the point of view of the writers, the best Japanese (J) word processing package should be the one that provides thoroughly automatic kana-to-kanji conversion at the time of kana/romaji input. For example, the latest MS-IME feature 'Auto correct' functions that can deal with predictable input mistakes such as adding a missing letter in geminate vowels and consonants, and automatically correcting a wrong use of alternative hiragana letter regarding one-to-two sound-to-hiragana correspondence. However, thus far there is no current application that can offer: either (a) user-free automatic processing of kana-to-kanji conversion, or (b) a complete detection of users' mistakes in the final output. In other words, J-word processing packages embody the implicit assumption that users are always able to recognize the right kanji or kanji compounds among a list of homophones. Thus, the role of writer is still vital for detecting a semantic anomaly or a wrongly selected homophonic kanji at local and the global levels of working text, even when writing with the most sophisticated IME. A further chapter (Chapter 7) will investigate an issue of L2 kanji verification process at the sentence level.

In order to illustrate the role of the writer, this section will demonstrate the conversion process of a two-kanji compound 'investment' 投資 [toːɕi]. The writer starts with kana-to-kanji conversion by typing its pronunciation *tousi*. Then the writer hits the return key when the string of hiragana と (to) う (u) し (si) is displayed on the screen. Just then, the writer sees a small pre-edit window appearing in the document with a list of kanji compounds that are pronounced [toːɕi]:

1投資	2 透視	3 凍死	4 闘志	5 闘士

Here, the role of the writer is to recognize 投資 for 'investment', and then kanji conversion takes place when the writer hits the fix key. If the first choice is not the desired compound, the writer needs to scroll the list until the desired one is displayed. Moreover, the inability to check wrongly chosen kanji or kanji compounds automatically makes information processing of those homophones entirely dependent on the kanji recognition ability of each writer. It is clear that the writing processes of Japanese kana-kanji mixed text require much more frequent interaction with the computer than those of the English language. For L1 adult users, this interaction is merely a troublesome chore, but so far as L1 children and L2 learners of Japanese are concerned, this 'dialogue' type of man-machine interaction has a strong potential to become an intelligent educational tool.

2.2.2 *Romaji input method and L2 writers*

There may be two major advantages in choosing the romaji input method in L2 word-processing. First of all, we can use existing English keyboard and associated operation skills such as 'blind touch' with the QWERTY keyboard. Secondly, writing the pronunciation of Japanese words with native Roman characters would be easier and faster for EJ learners because the more effective cognitive processes of sound-letter conversion could be facilitated by the use of L1 script in the romaji input method. At present, we do not know whether this is the case. This question of the effects of L1 orthography (romaji) on writing in Japanese will be taken up later in Chapter 6 (Reading and spelling of Japanese words).

However, the important point to note here is that Japanese orthography does not have a unified romaji spelling, and there are three slightly different versions in current use - (a) Hepburn Style, (b) Japanese Style, and (c) the Directive (b) (*kunrei*) Style (see Appendix B). Umemura et al. (1995, 180) state that the most frequently featured version is the Directive Style in Japanese IME, but they seem to

have overlooked the fact that the J-word processing packages actually allow the use of three styles of romaji in one application, in order to process some orthographic inconsistencies and homophones that I have discussed In 2.1.1. For instance, し/si/ can be obtained from either the sequence *si* (the Directive Style) or *shi* (Hepburn Style). Similarly, ふ/hu/ and ち/ti/ can be realized via *hu* or *fu* and *ti* or *chi* respectively. But the phonetic [ʑi] is normally realized in three different ways in order to process two different letters じ and ぢ. The letter じ is obtained by hitting either *ji* (Hepburn Style) or *zi* (The Directive Style), while the letter ぢ is only realized via *di* (Japanese Style). Likewise, the phonetic [ʑɯ] is realized by hitting *zu* (the Directive Style) for the letter ず, but *du* (Japanese Style) for the letter づ. To summarise, the romaji input method requires a unique usage of romaji spelling (see Appendix C), though most of the usages accord with one of the three existing romaji spelling conventions within the Japanese orthography system. This should be an important point to stress when a J-word processing package is used in L2 digital classrooms.

Relevant studies on Japanese language

Researchers (Koda, 1987, 1988, 1989; Chikamatsu, 1996; Mori, 1998) in Cross-language transfer studies agree that readers with different L1 orthography backgrounds behave differently in the 'phonological inaccessibility' condition (i.e. the absence of phonological clues in the linguistic item). Their key finding is that readers with alphabetic language backgrounds such as Arabic, Spanish and English were impaired in handling phonologically inaccessible elements embedded in a context, while reading comprehension or reading times among Chinese and Japanese readers in the same condition were not affected, because they employed more flexible and different strategies in line with their L1 logographic language background. In this light, English-speaking readers of Japanese are predicted to experience difficulties with the phonological inaccessibility of a given kanji/kanji compound. Added to this, Koda's cross-language transfer study has reported evidence suggesting a positive language transfer. She claims that the closer the distance between L1 and L2 orthographies, the more efficient language transfer from L1 to L2 should be, and this is supported by Tamaoka and Menzel's study described below.

Tamaoka and Menzel (1994) looked into the belief among L2 Japanese teachers that the use of romaji has a negative effect on L2 pronunciation of Japanese because of possible effects of L1 phonological interference. Their subjects were three

alphabetic native speakers (German, French, and English), one logographic subject (Chinese) and three L1 subjects (Japanese). Contrary to the belief, they found L1 positive script effects among alphabetic subjects, who processed the given text twice as fast as those in kana script, with reading accuracy being equal or better, but the reverse results were found with Chinese and Japanese subjects. They concluded that subjects made an efficient application of their L1 knowledge such as GPC (grapheme-phoneme conversion) rules to their L2 processing of romaji. This finding suggests a positive implication to the L2 use of the romaji input for J-word processing discussed above.

2.2.3 The impact of computer technology on Japanese literacy

A large number of studies in the last two decades have investigated the role of computers in an English environment, including the use of word processing packages in foreign language teaching (Daiute, 1985; Haas, 1996; Kellogg, 1994; Lam & Pennington, 1993; Pennington, 1996; Sanders, 1994; Sewell, 1990). According to a survey (Pennington (1996)) on the effects of computers in second-language (L2) education, there is general agreement that English word-processing packages have an effect on the writing process and the written product, but they have yielded inconsistent results on the question of whether the effect is positive (e.g. Daiute (1985), Lam and Pennington (1993)) or negative (e.g. Dowling (1994), Sanders (1994)). However, very little research has been conducted so far on the overall effect of educational J-word processing (Goto, Fukasawa, & Hamada, 2001; Hatasa, 1997; Otake, 1987; Tsuchiya, 2001), partly because Japanese pedagogy only began using computer writing in the 1990s. Consequently, many important pedagogical and methodological issues are yet to be addressed in this field. Thus, we need further research to obtain a critical understanding of the impact of computer writing environment on the development of L2 Japanese literacy and its pedagogical implications.

Different skills for digital kanji writing

Since the advent of J-word processors in the late 1970s [13], the relationship between the writer and text has fundamentally changed. In particular, the spread of personal computers with word-processing packages undoubtedly has altered our notion of kanji formation skills. Traditionally, kanji learners had to remember the specific stroke order of a kanji to reproduce the proper shape of that kanji stroke by stroke. Memorizing the shape of kanji requires constant hand-writing practice, rehearsal and repetition. But now, virtually no 'kanji writing' process is involved in obtaining the actual shape of a kanji when writing with J-word processors, because they automatically display a list of kanji on the screen for the user to 'recognize' the visual representation of intended kanji / kanji compounds on the display. To summarise, unlike traditional stroke-by-stroke handwriting, digital writing does not require a physical kanji-formation process; instead J-word processing packages require writers to have a reading-based kanji production skill. The more we write kanji with the keyboard rather than with the pen, the more the role of reading will take a central place in the kanji writing process. Here, by 'reading' I mean knowing both whole word pronunciation and the visual configuration of a given kanji or kanji compound. This is an important point to stress, as every chapter will come back to this point with different research perspectives.

A new role of kana literacy in the computer age

L2 specific input problem   As has been shown, kanji writing with the keyboard requires the precise phonological information of an intended word as an initial input. Recent literature on L2 use of J-word processors (Hatasa, 1997; Tsuchiya, 2001) identified an L2 specific input problem regardless of the writer's L1 background - L2 writers often cannot type the lexical phonology of an intended word accurately. For instance, Tsuchiya (2001) investigated phonologically oriented input errors of L2 writers from around 30 different L1 backgrounds, though he admits that his study largely concentrated on data from learners of Japanese from Asian language backgrounds (Chinese and Korean data constitute of 39.7% of the whole sample). According to his report (Tsuchiya, 2001, p.375), while input errors were influenced by L1 phonological interference and kanji learning styles of individual writers, there

---

[13] Japanese became a 'keyboard language' much later in 1978, due to linguistic and technical problems. There were three major problems: (a) encoding method to handle several thousand kanji, (b) input method, and (c) output methods including the display system and printing (Kanda, 1986; Kimura, 1988; Mori & Yagihashi, 1989).

was no clear-cut difference among writers from different L1 backgrounds. This situation is different from the L2 acquisition of Japanese phonology, where different problem patterns usually arise depending on the mother tongue. Tsuchiya (2001)'s report did not offer a full account for why there was a discrepancy between his findings and previous Interlanguage phonology literature. One reason could be the lack of European language data. For example, the report includes data from only 2 English native writers, involving only 334 Japanese characters. Thus, Tsuchiya (2001, p.377) suggests that it is important to collect more data from European learners at beginning levels, who have never stayed in Japan previously. Another explanation may be that the above research concentrated on phonological factors only and disregarded other factors such as the writers' keyboard fluency and orthography (hiragana and kanji) knowledge, which can also be associated with writing errors in word-processed texts. Firstly, as I pointed out earlier, J-word processing packages simply assume that their users can constantly recognize the desired kanji or kanji compound among a list of homophones. Thus, the less kanji knowledge the learners have, the more problems arise in their visual recognition and the more errors in the end-product as a result of poor self-monitoring. The findings from the following chapter will provide further evidence to this line of research that deals with L2 word-processed spelling errors.

A change in kana literacy practice   Next, I would like to point out that kana literacy factors may play a major role in L2 J-word processing. The reason is that the role of kana literacy in J-word processing is almost completely different from traditional handwriting. As I demonstrated above, the utilization of hiragana is much heavier in J-word processing than in writing by hand, since J-word processing takes the hiragana string as the basic orthographic unit that is converted into the kanji-kana mixed text. On the other hand, a typical handwritten text includes alternate use of kanji (content words) and hiragana (grammatical particles), in which the hiragana's main function is rather specific only to grammatical items such as postpositional particles and inflectional endings. On the whole, the Japanese pedagogy literature does not pay much attention to kana literacy acquisition, and heavily concentrates on issues in the more troublesome kanji learning, mainly because kana letters themselves are easy to learn as they hold a rather straightforward one-to-one grapheme-to-mora correspondence. Thus, exclusive hiragana usage is only seen in pre-school children's texts in L1 orthography learning, as content words and verb and adjective stems are gradually replaced with kanji according to the learners' kanji acquisition stage. Similarly, we rarely see exclusive hiragana usage in L2 academic

settings thanks to the aid of a kanji pronunciation indicator 'furigana' written in small print sized hiragana letters that appear either below or above the kanji.

The point is that the design of J-word processing packages involves a heavier utilization of kana orthography, thus changing our traditional literacy practice. For Haas (1996), a powerful force of computer technology 'shapes' the writer's physical relationship to the writing activities. However, the designers of J-word processing packages seem to have taken for granted that this was not an important issue for L1 adult writers since their kana literacy is already at ceiling level. Nevertheless, we do not know whether the same is true for developing writers. To the best of my knowledge, the issue of hiragana literacy in this perspective has never been discussed in the L2 Japanese pedagogy fields. It is worthwhile examining the subject more closely. The following section 2.3.1 will discuss previous findings from L1 children's reading acquisition literature on the causal direction of the relationship between phonological awareness and literacy.

Effects of computer technology on writing

A positive impact on quality of kanji writing   This section will focus on the positive and the negative effects of J-word processing on native and non-native writers. J-word processing packages are supposed to free writers from the boredom of writing practice and the difficult cognitive kanji-formation task. Indeed, Osumi (1992, 18-22) provided evidence of positive effects, where 12-year old grade-six Japanese children wrote more kanji with fewer errors on a Japanese word processor. He further reported (1992: 60-68) that there was not much difference in quantity of kanji between the pen-pencil group and the word processor group, but that the hand-writing group used more kanji from grade-1, while the word processor group wrote more difficult kanji, learnt in higher grades, or even some kanji that were beyond the elementary school learning items. Similarly, Jim, Kabashima, and Murakami (1993, p.138) report that L1 adult writers showed a tendency to write configurationally complicated kanji, which they would not normally bother to write, or could not write, by hand, when they used J-word processors.

In short, these reports provide positive evidence that the computer-writing environment certainly liberates L1 users from the tedious kanji retrieval task at the lower-levels of writing (i.e. spelling or kanji formations). Their findings coincide with findings from the L1 and L2 English literature (Bridwell-Bowles, Johnson, & Brehe, 1987; Daiute, 1985; Dunkel, 1991; Lam & Pennington, 1993; Sewell, 1990;

Susser, 1998). For example, Daiute (1985), who researches on the development of children's writing ability, claims that writing on a word processor is more like talking than writing, and word processors stimulated motivation for L1 12-year old American children. They wrote more and faster, because the computer not only relieves physical, psychological and cognitive constraints, but also gives children perfect-looking products. A typical conclusion drawn by these researchers is that a digital writing environment encourages the developing writers to concentrate on higher-level writing processes such as planning and editing, by providing them with lower-level skills such as spelling and formatting. As a result, they claim that word-processing packages facilitate the cognitive growth of the children and positive writing attitudes.

L1 writer's cognitive loss in digital kanji writing It should be noted that not all previous studies on word-processing provided positive findings. A serious claim is that computers and word-processing packages are detrimental for previously acquired literacy because the writing technology damages our mental representations of orthography (Sanders, 1994). A well-known dilemma among L1 Japanese computer users points to this issue. They often complaint that the more you write with the keyboard, the less you can recall and write the precise shape of a desired kanji when writing by hand. In other words, native writers believe that the daily use of J-word processing packages is somewhat related to a memory loss of once acquired kanji shapes, or the deterioration of hand-writing skill. According to Sato, Sugiyama, Furuhata, and Kobayashi (1996), one's kanji proficiency is socially more valued than before, since the double-edged nature of word processing - helpful, but involving the possible loss of an automatized cognitive skill (i.e. failure to recall the exact shape of an intended kanji), has been gradually known among users of J-word processors, which paradoxically makes people aware of the importance of their hand-writing skill and reading ability of kanji.

L2 writer's kanji writing strategy Within the CALL (computer-assisted language learning) literature, word processing packages are often referred to as open-ended tools, which generally liberate users from several types of workload, and encourage users to have a holistic approach toward their working process rather than locally handling each routine task (Underwood & Underwood, 1990; Sewell, 1990). Certainly, I often observe that L2 beginning users over-rely on electronically available bilingual dictionaries in our writing class at the computer class. For example, a second year Edinburgh student wrote in his course evaluation questionnaire that the supportive power of the bilingual on-line dictionary was so strong that he

could not read many kanji nor remember the meaning of some kanji in his essay that he himself had written not long previously. What his remark means is that writing with JWP gives him a prompt display of a list of suggested Japanese items with their English meanings via English input, and then the highlighted/recognized item will be automatically transferred to the main text by hitting one transfer command, without involving any element of the writing process of Japanese. Although this way of finding the desired kanji or katakana word makes the writing speed impossibly slow, the above student found this function helpful enough to compensate for his weakness in kanji knowledge. In sum, it was not a loss of existing knowledge or skills, but this example illustrates that the aid of an extended electronic lexicon could create an illusion of much higher performance level of kanji writing ability. If this example was not a one-off case, then written products in L2 digital texts may include superficially high quality kanji due to a new on-line writing skill (e.g. cut & paste). The following chapter includes written errors based on an inappropriate use of an on-line dictionary, and we shall see how such errors may affect readers' sentence interpretation.

2.2.4 Specifications of JWP

All of the relevant L2 data in the following chapter (Error Analysis) was written using a freeware J-word processor called JWP [14] that has been used in the Japanese computer writing class at the University of Edinburgh since 1996. Some researchers (e.g. Haas (1996)) working on effects of computer technology on writing claim that the design of a given word processing package can have substantial impacts on the writing process and its products. As I have no disagreement on this point, this section aims to describe the main characteristics of JWP.

Like other word-processing packages, JWP shares the basic principle of the kana-to-kanji conversion method, and it comes with an English-Japanese bilingual dictionary and a kanji look up dictionary. JWP can read and write Japanese language documents including Japanese email encoded in the major standards such as JIS (Japanese Industrial Standard) and Shift-JIS encoding, but it has some limitations, as the main priority of the creator of JWP was to give the users basic Japanese word-processing functions (Chung, 1995). As a result, font-control functions (e.g.

---

[14]PC users can download JWP from ftp.cc.monash.edu.au/pub/nihongo/ and this works with Japanese text without the need to install a Japanese language operating system. The basic architecture of JWP is very similar to widely used Word for Windows, so that it is easy to use even for the first-time novice writers.

bold and italics) and editing functions such as hiragana spelling checker are not supported by JWP. For this practical reason, positive or negative impacts associated with the Auto correction function will not be mentioned further in the following chapters. Similarly, the treatment of the inflectional endings of verbs is a little clumsy with JWP in comparison with a more efficient commercial IME. Furthermore, kana-to-kanji conversion in JWP normally takes place at the word level, rather than the more efficient sentence level conversion. In short, writing with JWP is more time consuming than writing with other commercial applications due to the less effective kana-to-kanji conversion. These limitations above indicate that JWP is not the most suitable writing tool for highly skilled writers. Nevertheless, I have observed through my teaching that this particular word-processor provides a considerably truthful reflection of spelling problems among developing L2 writers. Given this initial observation, the core of all questions in the further studies will be based on the general question of what factors play a role in L2 spelling difficulties when writing with the keyboard. Central to this issue is L2 acquisition of reading and writing. In particular, the following chapters will investigate how linguistic and perceptual difficulties have impacts on lexical reading and writing processes of EJ learners of Japanese.

2.3 Theories of speech and visual perception

2.3.1 Phonological Awareness

Phonological awareness and its relation to reading and writing have been active areas of research over the last 30 years. In that time, a number of issues have been investigated in L1 developmental research (Liberman, Shankweiler, Fischer, & Carter, 1974; Bradley & Bryant, 1985; Treiman, 1985; Mann, Tobin, & Wilson, 1987; Cunningham, 1990; Cunningham & Stanovich, 1990), L1 adult studies (Burt & Butterworth, 1996), bilingual research, Cross-linguistic studies (Cutler & Otake, 1994; Leong, 1991; Mann, 1986a), and L2 research. It is generally agreed that phonological skills and phonological awareness play an important role in alphabetic literacy (i.e. reading and spelling).

A traditional view in the literature sees phonological awareness as a metalinguistic skill, the application of which involves conscious reflection (e.g. (Mann, 1991; Liberman, 1971)). For example, Liberman (1971) considered phonemic segmentation problems as a potential cause of English speaking children's reading disabilities

such as dyslexia. The proposed reason was that English orthography represents its spoken forms in terms of phoneme-seized units. In this view, low levels of phonemic awareness are seen as the barrier to the acquisition of written language. Researchers have developed several means to measure the phonemic awareness of children. For instance, the relation between reading ability and phonological awareness can be tapped by tasks that require children to count the number of phonemes in spoken words or add or delete phonemes to words (e.g. Mann (1986b)). This tapping method to assess developing learner's phonological awareness is widely used in the literature, including L1 Japanese developmental studies (Inagaki, Hatano, and Otake (2000)). Although the tapping method has been used mainly in the L1 studies, I judge that an adopted technique should be practical to measure L2 phonological awareness of adult learners of Japanese. Accordingly, Chapter 5 on moraic awareness will employ a modified tapping method, which I believe is suitable for the measurement of L2 phonological awareness in Japanese among EJ learners.

In contrast to the view of phonological awareness as a metalinguistic skill, Treiman and others see it as a reflection of internal phonological representations (Treiman, Cassar, & Zukowski, 1994; Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995). This line of research provided an extensive study of invented spellings in English, and demonstrated children's reliance on phonology in phonologically accurate misspellings. For example, researchers such as Treiman and Zukowski claim that children's spelling errors such as "bet" for *bent* and "mik" for *milk* should not be considered solely as low level awareness of /n/ or /l/, but as a result of "mismatch between children's phonemic systems and those embodied in conventional English spelling". Ample empirical evidence from Treiman's studies pointed out that children's phonemic representations are not always the same as those assumed in the conventional writing system. This line of literature convincingly suggests that the structure of spoken language would provide a key foundation for studies of learners' literacy development. I favor this view, hence, findings and implications from this line of research will be extended to our L2 Japanese acquisition context in further chapters (Chapters 5 and 6).

Kana literacy and L1 moraic awareness research in Japanese

In 2.2.3 above, I pointed out that the field of L2 Japanese pedagogy lacks literature on hiragana literacy. At this point, I would like to mention that J-word processors clearly show spontaneous input errors as an unconverted string of hiragana letters.

What has to be noticed is that such hiragana errors are frequently seen in L2 JWP documents, and are hence noticeable. Is this simply because the design of J-word processing packages has made a change in traditional kana practice (i.e. heavier utilization of kana orthography)? If not, are these errors specifically created by a given word processing package JWP? In any case, a key question is about what causes this particular writing error, which is only associated with digital written products. Similarly, why do L2 learners seem to struggle when writing with hiragana-based J-word processors, although hiragana letters are generally considered as easy learning linguistic items?

While further studies aim to answer these questions, my working hypothesis here is that the knowledge of letter sounds alone did not equip EJ writers to encode the intended Japanese words perfectly. Recent studies in the L1 literacy developmental literature favor this view (Ball & Blachman, 1988; Dairoku, 1995; Inagaki et al., 2000). Their main aim was to explore young children's moraic awareness or moraic segmentation strategies, especially the effect of kana literacy on speech segmentation. This line of research provided evidence that children's phonological awareness plays a large part in learning to read. If the same can be said for the development of spelling skills, then an important determinant of spelling, like reading, should be moraic awareness. Support for this assumption comes from a wealth of evidence in English literacy development studies, which have established a strong link between children's knowledge of spoken language and their acquisition of spelling (Ehri, 1992; Goswami & Bryant, 1990; Muter & Snowling, 1997; Read, 1981; Treiman, 1993; Treiman et al., 1994). A fuller discussion on the relationship between reading and spelling will be followed in Chapter 6.

2.3.2 The syllable and the mora in Japanese

Treiman (1991) points out that phonological awareness refers to the awareness of several phonological units of the spoken language, including syllables, intrasyllabic units (i.e. onset and rime in English), phonemes, and phones. In other words, "Phonological awareness is not a unitary skill" (Treiman, 1991, p.160). Treiman rejected the traditional assumption that syllables are linear strings of phonemes, and put forward the notion of intrasyllabic units where syllables incorporate an intermediate units larger than the phoneme but smaller than the syllable, and Treiman's experiments have demonstrated that native speakers of English tend to segment the speech sounds at the onset-rhyme boundary (Treiman & Zukowski,

1988; Treiman, 1991, 1992). On this matter, most phonologisits would agree that any language has a level of prosodic structure below the syllable, and this can be seen as the moraic level in the case of Japanese (Kubozono, 1989, 1996; Ujihira & Kubozono, 1994). According to Kubozono (1996, p.85), "the nuclear vowel is combined with a preceding consonant to form a constituent called 'mora', while the postvocalic consonant or a post-nuclear vowel (i.e. the second half of a long vowel or a diphthong) forms another mora". Kubozono (1996) demonstrated a cross-linguistic difference in terms of preferred speech segmentation patterns between English and Japanese, by replicating Treiman (1986)'s blending experiment in which Japanese subjects were asked to blend two monosyllabic words into one blend form (e.g. big / rat). Interestingly, different blending patterns emerged for English and Japanese speakers, with the same set of materials. On one hand, English speakers in Treiman's study preferred to split the given word between the onset and the peak as C/VC (e.g. b(ig) / (r)at → bat); on the other hand, Japanese speakers preferred to segment the word at the mora (peak-coda) boundary as CV/C (e.g. bi(g) / (ra)t → bit).

While the above experimental evidence from Japanese suggests that morae are sub-syllabic units, it must be noted that morae in Japanese often overlap with syllables. As Kubozono (1989, p.250), points out, "syllables in Japanese generally consist of not more than two morae", and they can be defined phonologically as 'long syllables' (bimoraic syllables) and 'short syllables' (monomoraic syllables). For example, 'fujiyama' (*Mt. Fuji*) consists of four CV-type monomoraic syllables fu-ji-ya-ma, and constitutes a 4-morae 4-syllable word, whereas its synonym 'fujisan' consists of 3-syllables fu-ji-san that can be further divided into 4-mora units fu-ji-sa-N, since the word contains one bimoraic syllable CVN (i.e. 'san'). According to Otake et al. (1993), the former CV-type monomoraic syllables, in which mora and syllable coincide, constitute more than 70% of morae in Japanese speech corpora. In Japanese, long syllables in which V, CV, or CjV are followed by a postvocalic consonant N (nasal coda), Q (the first half of a geminate consonant), or a post-nuclear vowel R (the second half of a long vowel), are called *tokushuhaku* (special syllable). It is these special syllables (non-syllabic mora) that have been the focus of attention within Japanese pedagogy and Japanese second language acquisition, and especially in the Interlanguage phonology literature (Han, 1992; Hirata, 1999; Muraki & Nakaoka, 1990; Nagai, 1997; Oguma, 2001; Sukegawa, 1993; Sugito, 1989; Toda, 1998; Uchida, 1993). Relevant to this point is Kubozono's following remark Kubozono (1996, p.90):

The fundamental difference between mora-timed and syllable-timed languages lies in the durational difference between CVs and CVCs (or CVVs) and its function in the phonological system. In mora-timed languages the two types of syllable form distinct classes where CVCs and CVVs take twice as long time as CVs (in a psychological sense, if not in an absolute physical sense). In syllable-timed languages, on the other hand, the presence or absence of post-nuclear elements make virtually no difference and, hence, the two types of syllable play essentially the same role in the temporal organization of speech.

In other words, the control of the temporal organization of phonological durational contrast is essential for successful communication in Japanese. If learners of Japanese fail to control the expected mora timing with CVCs or CVVs, not only will their pronunciation end up with a foreign accent, but also, and more importantly, miscommunication may occur, since differences in vowel duration and consonant closure duration are contrastive for Japanese words. For example, the meaning of 'ku-ro' (*black*), 'ku-u-ro'(*air route*), and 'ku-ro-o' (*hardship*) are contrasted by the vowel duration (i.e. one or two morae) of each vowel. It is interesting to bear in mind that our EJ word-processed spelling errors, especially kana errors, have pointed to this problem. This would indicate that non-syllabic morae are also at work in L2 spelling difficulties in addition to the previous findings from Interlanguage phonology literature. Further chapters will look into this possibility in more detail.

L1-specific listening strategy

Recent psycholinguistic studies on segmentation procedures for speech sounds have been involved in an on-going argument about whether a basic segmentation procedure in a given language is based on any particular unit of perception, namely syllables, or it is determined by the language rhythm of each language (Bradley, Sanchez-Casas, & Gracia-Albea, 1993; Cutler, Mehler, Norris, & Segui, 1992; Cutler & Otake, 1994; Mehler, Dommergues, Frauenfelder, & Segui, 1981). Japanese is widely known as a 'mora-timed' language, in which the mora serves as a basic unit of temporal organization (Kubozono, 1989). In contrast, English has been said to be a 'stress-timed' language, in which the unit of rhythm is the stress unit. (Mehler et al., 1981) put forward the idea that the syllable is a universal segmentation unit,

while (Cutler et al., 1992) claim that speech segmentation is language-specific, thus the syllable-based segmentation would not play a role for English listeners' perception of French, because their stress-timed segmentation procedure is robust in their listening to foreign (in this case, French) speech sounds.

In the light of the language-specific hypothesis, Japanese speakers employ a mora-based listening strategy and segment speech at mora boundaries, while English speakers equipped with a stress-timed procedure segment continuous speech signals at the onset of strong syllables (Cutler et al., 1992; Otake et al., 1993; Otake, Hatano, & Yoneyama, 1996). Under this view, we would expect English monolingual speakers to apply their language specific stress-timed listening procedures to Japanese speech sounds, which do not share with English the property of lexical stress, regardless of their relative applicability. In normal listening situations, monolinguals do not have lexical access to a given foreign word, and they have no experience of comprehension problems in which their native listening strategy interferes with recognition processes. This is why monolinguals are insensitive to inadequacy of their L1-specific listening strategy for foreign words (Cutler et al., 1992).

However, questions still remain unsettled regarding the structural differences between Japanese and English, as well as the robustness of L1-specific listening effects on foreign language monitoring performance. Firstly, how can an English specific stress-based segmentation procedure be applied when English listeners are provided with Japanese language that holds a non-stress-based rhythmic structure? So far we know that under experimental conditions (Otake et al., 1996), Japanese listeners exploited their native mora-based strategy to French and Spanish, but not to English sounds. Otake et al. (1996, 198) accounted for this finding as a mismatch effect - Japanese listeners used mora-based strategy for English input as well, but because a 'stressed syllable is considerably lengthened in English', their expected durational value of CV as a single mora mismatched with the English stimuli, so that the mora-based segmentation strategy offered no opportunity for it to be utilized. Similar claims were made in Kondo (1998)'s production study, in which British beginning learners of Japanese were assumed to have transferred their English durational rule onto the production of Japanese words despite the fact that Japanese does not share with English the contrast between strong and weak syllables. According to their L1 stress-based segmentation procedure, an accented vowel was lengthened or the duration of other syllables was reduced before the accented vowel, which resulted in their failure to control durational contrast of Japanese.

The next question is the extent to which this L1-specific listening strategy is effective among English learners of Japanese, who are neither monolingual nor perfect bilingual, at different proficiency levels when they process Japanese input. Surely, we could reasonably suppose that monolinguals and bilinguals do not use identical listening strategies for foreign language monitoring. For example, Dupoux, Kakehi, Hirose, Pallier, and Mehler (1999)'s cross-linguistic (French and Japanese) perceptual study on the role of phonotactics at the perceptual level provides evidence that listeners use many different cues, including phonotactic information, for segmentation. Dupoux et al. (1999) were concerned with Japanese devoicing environments with [u], for which they digitally created non-word stimuli with 6 levels (from zero vowel to the original vowel), and asked their participants to make a judgment about whether there was a [u] vowel in the word. Dupoux et al. (1999) reported that only Japanese listeners predominantly hear illusory epenthetic vowels within consonant clusters even when the vowel was totally removed. Similarly, some English-speaking learners of Japanese in Toda (1997)'s production study applied their L1 specific segmentation strategy to a given Japanese word 'rikka' *first day of summer*, which consists of a CVCCV structure, and they interpreted the word as if it were two words CVC#CV 'rik-ka' similar to 'cat tail' and 'rock cake' in English. Although this strategy is illegal in Japanese as syllable phonotactics do not allow Q (the first consonant in a geminate cluster) to appear at word-final position (i.e. CVQ#), their use of this strategy successfully produced mora timing involving the phonological durational contrast for geminate consonants. Thus, Toda (1997) interprets this as positive L1 transfer.

Returning to the language-specific rhythm-based hypothesis, Cutler and Otake (1994, p.842) argue that bilinguals can inhibit ('switch-off') their rhythm-based listening strategies and employ 'generally available' segmentation procedures. This means that bilinguals exploit a language-specific strategy in their dominant language, but they 'switch-off' their rhythm-based strategy for segmenting non-dominant language sounds, and apply a 'generally available procedure' instead (Cutler et al., 1992, p.399). For example, French-English bilinguals who use syllable-based segmentation for French do not apply it when they are presented with non-dominant English sounds. Seen in this light, bilinguals can suppress their native language specific segmentation when the situation demands them to listen to a foreign word, but monolinguals cannot. Likewise, in a production study of Kondo (1998, p.39), fluent British speakers of Japanese managed to control durational contrast in Japanese, because they abandoned their stress-based listening strategy, and used an L2 mora-based strategy for their production of Japanese words. However, this clear-cut

strategy may hold only among perfect or balanced bilinguals, since other studies (Bradley et al., 1993; Kearns, 1994) report that the choice between the use of language-specific segmentation strategies and more general sub-lexical segment-to-segment strategies was not seen among less perfect bilinguals, including French learners of English. Moreover, Cutler and Otake (1994, p.842), report that their bilingual listeners demonstrated native levels of production and comprehension skills in both languages regardless of the segmenting procedure they commanded. All of this amounts to saying that there is no conclusive view about the effect of L1-specific listening strategies in foreign language learning environments, and the dynamic or developmental nature of Interlanguage phonology at different stages of acquisition could be an important factor when we investigate the effects of L1-specific language listening strategies for the processing of foreign language sounds. We shall take these previous findings into account for the development of the experimental materials for Chapter 5.

2.3.3 Visual word recognition and reading kanji

Models of reading

A good place to start is the dual-route model. In this approach, there are two possible processing routes for recognizing visually presented words:[15] (a) the semantic route (a direct access or lexical route), and (b) the grapheme-phoneme conversion (GPC) route (non-lexical or sublexical route). A key feature of the general dual-route models is the assumption that separate processing mechanisms are required for the pronunciation of irregular or exception words and nonwords. For instance, the direct access route is used for words such as 'pint' with irregular spelling-sound correspondence, whereas nonwords like 'zaid' cannot have direct access, because they do not have lexical entries, but pronunciations can be assembled by means of the GPC route (Coltheart, Curtis, Atkins, & Haller, 1993).

In this model, normal readers are supposed to use both the direct route and the GPC route, and a race process begins when the word is seen. Generally, however, the direct route is considered to be much faster. Earlier, the two routes were assumed to be independent of each other, but some evidence has shown that they are less independent than was previously suggested (Glushko, 1979; Seidenberg, Waters, Barnes, & Tanenhaus, 1984). For instance, a study reported by Seidenberg

---

[15]There are also multi-route models within this framework.

et al. (1984) found that the *regularity effect*[16] interacts with frequency, namely, this regularity effect was found with low frequency words but not with high-frequency words. This suggests that we name familiar words equally fast regardless of their GPC regularity. This further implies that we will employ GPC rules when we try to read unfamiliar or low-frequency words, due to the rather slower process of the direct route. To sum up, Coltheart et al. (1993) argue that their model can provide answers to the reading phenomena such as how skilled readers read nonwords and exception words aloud, and how acquired and developmental dyslexia arise. However, we must note that these findings were obtained from experimental studies done with English words. As I have demonstrated with example 1 in page 6, Japanese orthography comprises two primary scripts, namely, syllabic kana, and logographic kanji. Many Japanese Psycholinguistic studies have addressed questions about the nature of kana/kanji processing, such as whether or not kana and kanji processing employ different mechanisms (Sasanuma, 1975; Nomura, 1981; Henderson, 1982; Hatta & Konda, 1992; Nagae, 1992; Flaherty, 1993; Leong & Tamaoka, 1995). The relationship between Japanese orthography and phonology in kana/kanji processing appears obviously more complex than in English words.

Before we come to kanji reading research in detail, we shall touch upon Seidenberg and McClelland (1989)'s parallel distributed processing (PDP) connectionist model of visual word recognition and reading. We shall not go into detail, since our main purpose here is to grasp the key processing mechanism within this model. The fundamental property of their approach is that the PDP model needs only a single route between orthography and phonology. The basic architecture of this model is a computation mechanism called back-propagation[17] net, which can learn to pronounce all types of letter strings (words, nonwords, and irregular words) through repeated training or experience with the spelling-sound correspondences found in the set of words. The simulation process is based on three units: (a) orthographic units (the input), (b) the hidden units, and (c) phonological units (the output), where the network can learn to connect the input (spelling patterns) with the output (phonological codes) in terms of the weights on connections between units[18]. The connectionist model differs from the dual-route model in that it does not assume a lexicon, because words are represented by a pattern of activation over the

---

[16]It takes more time to read GPC irregular words than regular ones.

[17]'back-propagation' is a learning algorithm.

[18]All orthographic and hidden units are connected, similarly all hidden and phonological units are connected. In addition, the orthographic units receive feedback from the hidden units.

hidden units, hence, no *lexical access*[19] stage is involved. The key pedagogical im-
plication of the connectionist view may be that the PDP framework is a learning
model. Because of this learning mechanism, Seidenberg and McClelland claim that
the model can provide a good account of several facts about normal and abnormal
reading. For instance, "differences between readers in terms of word recognition
skills": e.g. skilled vs. unskilled readers, developmental and acquired dyslexias,
and "transitions from beginning to skilled reading" or the acquisition process (Sei-
denberg & McClelland, 1989, 523-546). The PDP model concerns how the readers'
knowledge of the links among orthographic, phonological, and semantic informa-
tion is represented, how they are computed, and how the computed information are
used for the different reading tasks. According to Seidenberg and McClelland (1989,
560), *frequency* affects these computations; to put it another way, frequency as "the
reader's experience in reading, hearing, and pronouncing words" has an important
role in reading. Although PDP models may have interesting implications for kanji
learning, they will not be discussed any further in this thesis.

Turning now to kanji reading, the key question should be that of which type of infor-
mation - visual, phonological and meaning, exerts the major influence in kanji word
recognition. Two different research findings are reported in Sakuma et al. (1998,
77). According to one view, skilled word recognition does not require phonological
information, and the meaning is activated directly by visual information (the direct
access view). On the other hand, a neuropsychological study of Japanese Alzheimer
patients with dementia gave evidence against this direct access view of kanji, show-
ing that the pronunciation of kanji words can be achieved directly by orthography
without taking a semantic path. Unlike the dual-route model, some researchers put
forward a strong phonological model in kanji recognition by claiming that phonolog-
ical coding is an automatic process and an essential component in word recognition
regardless of orthographic types (Perfetti & Zhang, 1991; Wydell, Patterson, &
Humphreys, 1993). In short, theorists currently agree that both phonological and
semantic information in kanji appear to be activated in kanji recognition, but there
has been some theoretical controversy about the issue of when phonological infor-
mation has its effect.

---

[19]'lexical access': i.e. entering the lexicon with its store of detailed information about words
(Eysenck & Keane, 2000, 533).

Phonological models of kanji reading

The traditional view of reading kanji favors the direct access view; that is, visual information activates the meaning of the kanji directly, without the use of a phonetic path, and so phonological information plays very little part in processing kanji. On the contrary, phonological information is the primary route from orthography to meaning in phonological models. For instance, in the case of Chinese, Perfetti and Zhang (1991) provided counter evidence against the traditional view. They claim that the sounds of kanji were activated, even in silent reading, immediately at the level of word identification. For them, phonological processes play a significant role in comprehension and memory, and they put forward a strong phonological activation view of Chinese lexical access. However, as Kess and Miyamoto (1999, 34) point out, phonological activation in Japanese lexical access is not as straight forward as in Chinese, because Japanese kanji are far more phonologically ambiguous than kanji characters in Chinese (hanji). This is a result of the possibility of multiple readings for a given kanji, and the existence of a large number of kanji homophones. This issue will be discussed later in relation to a parallel-access view, but here I would like to focus attention on some strong phonological models of Japanese kanji reading.

Horodeck's classic study

Horodeck (1987) questioned the traditional view of kanji reading and posed a question of whether kanji are really ideographs that trigger only meaning when Japanese is read. He conducted an L1 reading study with 219 native adult speakers of Japanese. Horodeck provided 60 newspaper headlines, in which homophonic and non-homophonic kanji errors were embedded. The kanji errors were created with regards to their form, meaning and sound. For instance, homophonic and non-homophonic errors were visually very similar but distinct in meaning to the correct kanji. He also inserted 10 % "dummies" (homophonic on-compound with wrong form and meaning): e.g. 'stairs' 階段/kaidaɴ/ for 'conference' 会談/kaidaɴ/, to check whether L1 readers 'really' read for meaning. L1 readers were given no more than five minutes to detect and mark the kanji errors in those 60 headlines. The result showed that about 53 % of L1 readers detected every "dummy" type of heterographic on-compound homophone, but they regularly failed to detect homophones which were visually very similar to the correct kanji. The important finding

was that L1 Japanese adult readers ignored inappropriateness of meaning or pro-
nunciation when there was a close visual resemblance between the errors and the
correct kanji, and failed to detect homophonic errors significantly more often than
non-homophonic errors. Although the first part of his finding seems to suggest some
active role for visual information rather than phonology, Horodeck concluded that
kanji are not ideographs for L1 readers, but they function as symbols for sound
when Japanese people read and write.

Recent studies have further extended Horodeck's conclusion. For instance, Mat-
sunaga (1995) conducted an eye-movement experiment on phonological coding in
reading kanji. Matsunaga's aim was to test the validity of Horodeck's study, es-
pecially about whether he measured "true reading" or "error monitoring". The
eye-movement patterns showed that the non-homophonic errors had significantly
longer fixations than the homophonic errors, and more non-homophonic errors than
homophonic errors were detected by L1 adult readers. In other words, L1 readers
were less distracted by homophonic errors than sound dissimilar non-homophonic
errors. The finding provided strong evidence to support Horodeck's study. Mat-
sunaga confirmed that Horodeck's study was not error monitoring, but true reading
for meaning.

The effect of orthographic similarity

We noted a little earlier that Japanese kanji are phonologically opaque for a large
number of homophonous kanji. In the previous research on English word recogni-
tion, researchers found a *homophony effect*.[20] The key issue here is whether this
effect was found as a result of phonological or visual similarity. Apart from this, re-
searchers seem to agree that (a) not all nonwords are equal, and (b) something else
besides phonological recording is involved. Many researchers in the kanji reading
literature have paid a great deal of attention, as we have already seen, to the issue
of the interactive relationship between orthographic similarity and phonological in-
formation. For the present, it may be useful to look more closely at the literature
on homophony effect. We shall first look at the literature on homophony effect with
English words.

Van Orden (1987) investigated the effect of homophony in a semantic category
judgment paradigm with English homophones. His subjects were asked to judge

---

[20]It takes less time to name or read aloud pseudo-homophones (e.g. *brane* for *brain*), but longer
to reject them as nonwords (see Harley (1995, 105-106)).

whether a presented target word (e.g., *tulip, rows,* or *robs*) belonged to the category of 'a flower' that they had just seen. In the target words, he inserted a non-flower category item *'rows'* that was a homophone of a visually similar real member of flower *'rose'*. The results indicated a strong homophony effect - there were more categorization errors to homophone foils like *'rows'* than non-homophonic controls such as *'robs'*. The effects were largest when the homophone target words were visually similar to the correct exemplar of the category name (the effect of visual similarity). In addition, this false-positive decisions (e.g., *'rows'* as a *'flower'*) were made even after the subjects spent longer time for *'rows'* than *'robs'*. He claimed that this evidence supported his strong phonological mediation model of word recognition, which views the primary role of bottom-up phonological procedure in orthography-to-meaning process. This claim has been challenged by Sakuma et al. (1998). We shall return to their study later in this section.

In the preceding section I discussed Horodeck (1987)'s study on L1 kanji reading. We notice that his results showed remarkable similarity to the study of Van Orden (1987): i.e. L1 Japanese adult readers failed to detect homophonic kanji targets significantly more often than non-homophonic kanji targets. However, a strong visual similarity effect was also reported in Horodeck. The following Japanese research all seems to agree that the similar homophony effect as well as the visual similarity effect should be found during kanji recognition (Wydell et al., 1993; Matsunaga, 1995; Sakuma et al., 1998). Wydell et al. (1993) pursued Van Orden (1987)'s study with Japanese kanji words. Their study found a very similar result, with effects of both visual similarity and homophony on semantic judgments. In addition, the visual similarity effect accelerated the homophony effect further. In short, these studies suggest a closer interaction of phonological and visual information in kanji reading. A more recent study by Sakuma et al. (1998) came up with even stronger evidence in favour of the effect of visual similarity. They claim that the primary information source during print to meaning processing is orthography. In other words, their claim suggested a major effect of the visual information in kanji reading. This is interesting, because their conclusion does not agree with a claim put forward by Van Orden (1987). We shall consider this matter further.

In Sakuma et al. (1998), they confirmed that their study was consistent with the previous findings that were found in Van Orden (1987) and Wydell et al. (1993). However, they reported that their third study with a pattern masking condition [21]

---

[21]Pattern masking or feature masking involves displaying the stimulus for a short period, and subsequently presenting an unrelated "mask" stimulus. This interferes with the word recognition process so that subjects are unaware of the stimulus.

showed a strong impact of orthography. They point out that English and Japanese orthographies establish clear-cut differences in the early processes of visual word recognition. In the case of English experiments with a pattern masking paradigm, phonological information is activated quickly enough to influence semantic decisions. On the contrary, their similar experiment in Japanese suggested the partial activation of phonological information, which was not early or full enough to influence semantic decisions under masking conditions. They concluded that orthography exerts the major influence on the activation of meaning for kanji words. Their interpretation was that their results were compatible with both the phonological recoding view and the parallel access view [22] of word recognition. In other words, their view suggests a close interaction between orthographic and phonological information in kanji reading. I agree with this point that the sound value of kanji has an important role in kanji reading, but it is not the all-or-none type of information indicated by some models described earlier. However, we should not overlook that these studies, except Horodeck (1987), was done with kanji or kanji words in isolation, which excludes information from contextual top-down processing. We may also need further research that will investigate the contextual effect on embedded kanji in the written texts. This will be investigated further in Chapter 7.

---

[22]Parallel access view of word recognition: both orthography and phonology contribute to the activation of the meaning of written words. See Sakuma et al. (1998, 85).

CHAPTER 3

ERROR ANALIYSIS OF L2 J-WORD PROCESSING

3.1 Motivation and research questions

The aim of this chapter is to identify the extent of the L2 spelling problems that can result from erroneous input into J-word processors.[1] In chapter 2, I have pointed out that the computer-writing environment has changed our notion of kanji formation skills and kana literacy practice, for which the writers need to apply 'reading' based writing skills. Although recent literature on L2 use of J-word processors identified an L2 specific input problem regarding inaccurate spelling of the lexical phonology of an intended word, previous research in this field has not provided us with a complete picture of L2 J-word processing. We first need to clearly pinpoint the problems faced by L2 users of J-word processors before we can discuss pedagogical practice such as computer writing instructions and learning environments. I thus set as my first task to ascertain whether I can make findings similar to those of previous research, with reference to a small spelling error corpus[2] collected from my own English-speaking students who are studying Japanese as a foreign language (hereafter EJ learners). In the following sections, we shall not treat the errors in the corpus as a random sample of all errors made, but to attempt an explanation for the errors that were recorded. The interest in the current study is how particular errors shed light on the underlying units of linguistic performance of L2 developing writers.

---

[1]This chapter is based on a journal paper by the author (Matsumoto-Sturt, 2003).

[2]According to Corder (1973), the study of learners' errors can be used to test the hypothesis similar to experiments.

I predicted that I would find common features among word-processed EJ-written errors, the pattern of which should provide evidence for the existence of spelling problems. These patterns may also suggest where the EJ writer's main cognitive burden lies. Here we are particularly interested in investigating the idea that the burden lies in the lower-level writing processes (i.e. lexical spelling), including the use of newly acquired digital orthographic skills in a complex domain made up of linguistic knowledge, visual and speech perception, and meta-cognition (error monitoring). The question that was therefore addressed in this study was "Does the use of J-word processor really liberate EJ writers from lower-level routine tasks? If not, to what extent do the identified spelling errors reflect the lower-level writing problems?"

3.2   Method

3.2.1   Establishing the corpus of learner spelling errors

Previous studies on Error Analysis

Error Analysis (EA) in second language research has a long tradition dating to the 1970s (Corder, 1974; Dagneaux, Denness, & Granger, 1998; Hatta et al., 1997; Frith, 1980). Especially, foreign language teaching specialists regard traditional EA as a window on the developing learners' Interlanguage. Although there is no doubt that EA provides useful information about error typologies made by L2 learners, EA's shortcomings have often been criticized (see Cook (1993), Ellis (1985)). A major limitation is that EA only gives a static picture of L2 learning, so that it is normally difficult to determine how a given error was made. In addition, there are some methodological limitations, such as the difficulty of establishing a set of well-defined error classifications (Nelson, 1980; Dagneaux et al., 1998). In this connection, Nelson (1980, p.476) comments that error frequency measures are often unreliable due to ill-defined error categories that are not mutually exclusive. Moreover, treatment of badly miss-spelled words, which normally include multiple errors, often involves a high degree of subjectivity. For this reason, Nelson (1980, p.477) points out that inclusion of an 'unrecognisable' category as a "safety net" is not good enough to produce reliable research outcomes, due to the difficulty of interpretating heavily distorted errors. Nelson (1980, p.477), thus, suggests that investigators should deal with multiple errors according to whether or not the error satisfies certain criteria, so that even a heavily distorted multiple error can be

treated similarly to a single error. For this reason, I have attempted to devise a detailed and sometimes redundant error classification system to deal with our L2 leaner data.

According to Granger (1998), the emergence of the computer corpus in the early 1990s has made it possible for automated linguistic analysis to be based on quantitative approach. Recently, language researchers and teachers have started to use DIY small corpora as a practical means of carrying out classroom-centred research (e.g. Ghadessy, Henry, and Roseberry (2001), Matsumoto-Sturt (2002a)). In this respect, computer-aided EA has an advantage over traditional EA in that the former can generate much more robust frequency counts, and enables automatic sorting of any number of errors to generate comprehensive lists of error typologies. In order to compile learner data in a machine-readable form, Dagneaux et al. (1998) recommends to follow strict design criteria that are along the lines of the general principles of corpus linguistics. For example, learner factors such as age, L1 background, learning stages, and learning situations (e.g. type of writing mediums, classroom SLA, etc.) should be considered. Thus, I have taken the following factors into account: learner type (English-speaking learners of the same age group), medium (digital writing with JWP), length, genre (essay writing), and learning context (task based class-room learning). In short, unlike early EA studies, our L2 data has been compiled electronically, thus quick and efficient manipulations of data can be expected. For example, a combination of the location of the error and each mora type can be searched within a short period of time.

3.2.2 Source of data

The errors listed in this study were all those found in the Japanese compositions of 43 (21 male and 22 female) university students[3], who enrolled on a elementary level 'Japanese 2A' course at the University of Edinburgh, between 1997 and 2000. Of this group, 37 students were native speaker of British English, 2 students were native speakers of American English, and the remaining 4 students were bilingual British English speakers with Thai (n=1), German (n=2), and Chinese (n=1). The average age was 21.3 (R : 19-27, SD=1.928), 90.7% (n=39) had never lived in Japan, and 81.4% (n=35) had 205 hours of learning Japanese at the same courses. Accordingly, the error corpus can be identified as a 'homogenous sample' in terms of

---

[3]Although there were more students in this course, only compositions written by English-speaking learners of Japanese were selected for the purpose of this study.

learners' age group, L1 background, and stage of development as recommended by Error Analysis procedure (Corder, 1974; Dagneaux et al., 1998). The given topics were 'a typical event of my country', 'a famous person in my country', 'famous arts or sports in my country'. Students chose one of these topics to write a word-processed composition in Japanese, all using the identical J-word processor JWP (see 2.2.4) during the class. Students were allowed to use both on-line and off-line dictionaries. The completed compositions, worth 15% of the course, were submitted after attending 18 one-hour per week computer writing classes.

3.2.3 Identification of errors

The error corpus was obtained by examining all of the misspelled Japanese words including hiragana, katakana, and kanji. In searching for errors I recorded any items that deviated from the spelling of the Shinmeikai Nihongo Accent dictionary except apparent grammatical formation errors. In addition, any deviations against the contemporary orthographic conventions, which are recommended by the Japanese governmental authority, were also recognized as kana errors. As for kanji written errors, homophone kanji errors are widely known as a typical L1 digital kanji error. However, unlike in handwritten kanji error corpora, no 'misspelled' kanji can be found in the digital written text in Japanese. Nevertheless, I suspect there may be another type of L2 specific techno-kanji error in addition to the homophone kanji errors that are familiar from L1 writers. Therefore, I scrutinized for non-existing kanji words (non-word), and homophone kanji errors in a given sentence as an erroneous kanji. For every error, I noted, if possible, the target word intended by the EJ writer. In all, 366 orthographic errors were collected. With an average text length of 56.4 lines, per composition, the average number of Japanese characters[4] was 1,693. Of the 366 errors, 13 involved an identical error or several repetitions of the same error by individual writers. After excluding such recurrent errors, the total number of errors was reduced to 340.

3.2.4 Classification of errors

Each word was assigned eight major codes: (1) orthography type, (2) word class, (3) semantic traceability, (4) homophones, (5) error type, (6) mora type, (7) location of

---

[4]Unlike English essays, Japanese essays are not counted by the number of words, but number of characters.

phonological error in the erroneous word, and (8) reading type. Both the 8 major categories and their subcategories were originally set for the present study. The error categories will each be discussed below.

Category 1: Orthographic type This category contains 5 subcategories: (1) hiragana error, (2) katakana error, (3) kanji error, (4) kanji plus hiragana error, and (5) other. While categories (1) to (3) were straightforward, categories (4) and (5) cover more complex errors. Category (4) was set for errors in inflectional endings, so that we can separately count segmentation errors related to okurigana (i.e. inflectional ending of a verb expressed by hiragana). Category (5) was established as a miscellaneous category. For instance, this category would include heavily distorted errors with several errors in a chunk, or a strange mixture of kanji and Arabic numerals.

Category 2: Word class 7 subcategories were created for this category: noun, verb, adjective, particle, adverb, conjunction, and unclassifiable.

Category 3: Semantic traceability This category separates semantically traceable errors from heavily distorted words.

Category 4: Homophones The error types in this category are: (1) homophone, (2) non-homophone, or (3) unclassifiable.

Category 5: Error types All errors are subcategorized as (1) additional, (2) omission, (3) substitution, (4) transposition, (5) blend, or (6) mix (more than one error), and (7) covers visual errors.

Category 6: Mora types Japanese allows a very restricted phonological inventory, and only 5 types of syllable CV, CCV, V, CVC (nasal coda and geminate consonant) are permitted (Kubozono, 1999). Since the notion of mora plays a crucial role in Japanese orthography, especially in the moraic writing systems hiragana and katakana, this category includes 6 essential morae: (1) short vowel (V), (2) long vowel (VV), (3) CV, (4) CjV, (5) geminate consonant (Q), and (6) nasal coda (N). In addition, 2 more subcategories (7) mix (more than one error), and (8) word-level (no phonological mistake is involved) were added.

Category 7: Location of input error in an erroneous word The location of the error in the erroneous word is to be identified by (1) the first mora, (2) mora

other than the first or last mora (henceforth, mid morae), or (3) the last mora. There are 2 more subcategories: (4) inflection, and (5) unclassifiable. Subcategory (4) was used especially in order to separate written errors, which were apparently caused by grammatical formation errors, from non-derivational written errors. Any errors regarding vowel durational contrasts will be classified according to criteria 1 and 2 below:

1. If vowel lengthening occurs at the first mora position, the error will be sub-classified as (1), though the actual written error itself exists at the second mora position.

2. If vowel shortening occurs at the second or later position except for last mora position in the erroneous word, the error will be sub-classified as (2).

Category 8: Reading types This category handles multiple readings of kanji in Japanese. The subcategories are (1) ji-on (Chinese reading), (2) ji-kun (Japanese reading), (3) mixed reading, and (4) other (e.g. words written by katakana).

3.3   Error Analyses

3.3.1   Distribution of errors by orthography type and word class

Tables 3.1 and 3.2 below present the obtained distribution of errors in each orthographic type and word class in the entire corpus:

Table 3.1: Number and Percentage of Errors (n=340) by Orthographic Types

Hiragana	120	(35.3%)
Katakana	51	(15.0%)
Kanji	51	(15.0%)
Kanji + hiragana	99	(29.1%)
Other	19	(5.6%)

Of the large number of hiragana errors (n=120) in Table 3.1, most are concerned with problems in kana-to-kanji conversion with the romaji input mode. Recall that the keyboard input system of J-word processors begins with getting hiragana letters on the screen by typing the whole word phonology of the intended word. We could consider this distribution pattern to show that EJ writers have phonological

Table 3.2: Number and Percentage of Errors (n=340) by Word Class

Noun	132	(38.8%)
Verb	125	(36.8%)
Adjective	26	(7.6%)
Particle	20	(5.9%)
Adverb	6	(1.8%)
Conjunction	3	(0.9%)
Unclassifiable	28	(8.2%)

input problems like those of L2 writers in the previous studies of Goto et al. (2001) and Tsuchiya (2001). We shall seek more tangible evidence in further analysis to discover whether this is the case.

3.3.2 Location of phonological input errors and the mora type

Japanese is a typical CV language, and simple CV type morae account for over 60% of all possible morae (Otake et al., 1993). Undoubtedly, this skewed distribution towards CV type morae in Japanese is reflected in my corpus as CV marked the most frequent error type in 3 subcategories; first mora (n=14), mid morae (n=25), and inflection (n=62) in Table 3.3 below.

Table 3.3: Mora type and the Location of Errors

	V (n=18)	CV (n=126)	VV (n=41)	Q (n=37)	N (n=8)	CCV (n=10)	Mix (n=56)	Word-level (n=44)
First mora(7.4%)	2	14	3	0	0	6	0	0
Mid morae (19.1%)	3	25	18	10	4	3	2	0
Last mora (6.5%)	1	3	12	1	4	0	0	1
Inflection (33.5%)	11	62	5	25	0	1	10	0
Unclassifiable(33.5%)	1	22	3	1	0	0	44	43

Actual samples taken from the error corpus are shown in 1) to 14) below. All examples of Japanese errors will appear in a box as shown in 1). Next, the whole word phonology and the meaning of each intended word are given in PM: e.g. PM-1 for 1. The Roman input of each error word is presented in the equivalent number with R[5]: e.g. R-1 for 1, where errors will be marked as follows:

---

[5]The segmentation point of each example was made according to the corresponding hiragana characters (morae) in the error word. Roman characters in R represent the typical writing input of Roman characters for J-word possessors, and do not always correspond to Romaji conventions in Japanese.

1. Underline marks the position of an error,

2. Framed items in R indicate substitution errors with two-to-one correspondence (i.e. two items share the same phonetic value) between alternative hiragana letters or homophone kanji,

3. 'ϕ' marks an empty mora (i.e. omitted mora),

4. Curly brackets show the surrounding environment of a given error,

5. A small dot between two items indicates a word boundary,

6. The intended word and the correct Roman input will be presented in parentheses next to the erroneous example in Japanese and R-examples respectively.

1) a. し [が] し, b. の [ひ] なやむ, c. そ [こ] ごろ, d. サンフランシ [ツ] コ, e. つ [ず] けました

R-1 a. si-ga-si (si-ka-si), b. no-hi-na-ya-mu (no-bi-na-ya-mu),
   c. so-ko-go-ro (so-no-ko-ro), d. sa-n-fu-ra-n-si-tu-ko (sa-n-fu-ra-n-si-su-ko),
   e. tu-[zu]-ke-ma-si-ta (tu-[du]-ke-ma-si-ta)

PM-1 a./sikasi/ 'but' b./nobinayamu/ 'failed to grow' c./sonokoro/ 'in those days'
   d./saNfuraNsisuko/ 'San Francisco' e./tuzukemasita/ 'continued'

Of 25 CV errors in the mid morae position 72% (n=18) are substation errors. Such examples are shown in 1). For example, 1a) and 1b) are examples of voiceless-voiced CV substitution. *sokogoro* in 1c) appears to be the result of semantic blending of two phonologically similar words *konogoro* (recently) and *sonokoro* (in those days). Example 1e) shows the mistaken choice of one of two hiragana characters that share the same phonetic value [zu]. Some examples of the next most frequent errors in the mid morae position regarding long vowels are shown in 2).

[h] 2) a. スバスター, b. [よこ] さん (洋子), c. [諸] 説 (小説) d. Tシャ [ー] ツ, e きょ [お] ぎじょう,

R-2 a. su-ϕ-pa-ϕ-su-ta-— (su-—-pa-—-su-ta-—)[6], b. yo-ϕ-ko-sa-n (yo-u-ko-sa-n),
   c. sho-ϕ-se-tu (sho-u-se-tu), d. thi-—-sya-—-tu ([thi-—]-sya-ϕ-tu),
   e. kyo-[o]-gi-jo-u (kyo-[u]-gi-jo-u)

PM-2 a. /suupaasutaa/ 'superstar' b. /yooko.san/ 'Yoko-san' c. /sjoosetu/ 'novel'
   d. T-/sjatu/ 'T-shirt' e. /kjoogijoo/ 'ground'

The first 3 examples (2a-c) are omission errors, while 2d) is an example of an insertion error. Again 2e) is an example of illegal hiragana usage similar to 1e) - the sequence of long vowel should be presented by hiragana う/u/, not with お/o/. Of 18 VV errors in this position, katakana errors (n=10) were concerned with omission

---

[6]The second vowel in VV sequence in katakana words is conventionally written with a long bar '—'.

(n=8) and addition (n=2) errors, while the remaining 8 errors (omission=3, substitution=4, mix=1) were all VV [oː] errors; especially all substitution errors were related to illegal hiragana usage between the right use of hiragana う/u/ and the wrong version お/o/.

At the last mora position, the most frequent errors were either addition (n=8) or omission (n=4) of the second sequence of VV as shown in 3).

3) a. 裏ばなし い, b. いっしょ う (一緒), c. 映画など う, d. 石 い, e. ディズニ

R-3 a. u-ra-ba-na-si-i̱ (u-ra-ba-na-si-φ), b. i-s-sho-u̱ (i-s-sho-φ), c. e-i-ga-na-do-u̱ (e-i-ga-na-do-φ), d. i-si-i̱ (i-si-φ), e. dhi-zu-ni-φ (dhi-zu-ni-—)

PM-3 a. /urabanasi/ 'inside story' b. /iQsjoo/ 'one's whole life' c. eiga /nado/ 'movies and so on' d. /isi/ 'stone' e. /dizunii/ 'Disney'

As a summary, the data in Table 3.3 are shown in graphical form in Figure 3.1 below, detailing the distribution of error locations in each of the eight subcategories.

Figure 3.1: Mora type and location of error
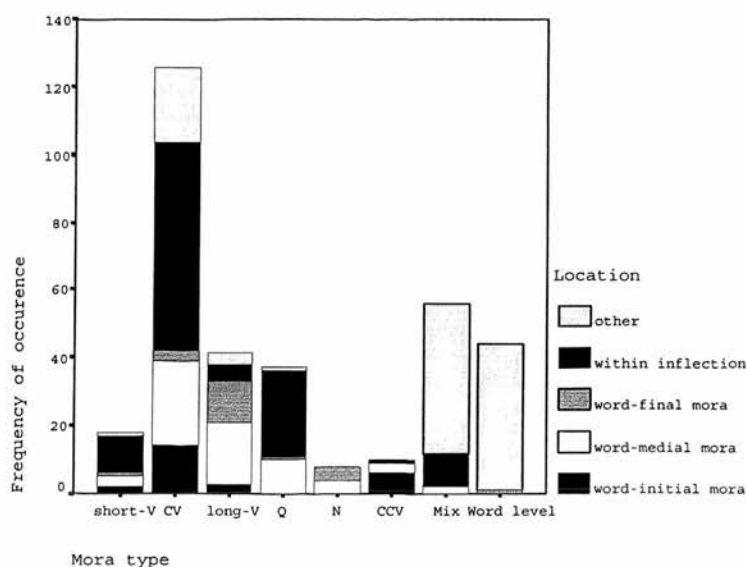
First, the most frequently occurring errors in the 'within inflection' category were CV substitution (n=29). Of these within-inflection CV substitution errors, 86.2% were verbs as shown in 4) below. Next, the second most frequently occurring error type in this category was CV omission (n=17) with about 71% (n=12) errors that were related to the wrong division of the kanji stem and the inflectional ending expressed by hiragana (i.e. okurigana) such as 行ました /i(ki).masita/ for the right

version 行きました /i.kimasita/ 'went'. Thirdly, substitution between the voiced and voiceless versions of CV morae occupied about 41% (12 out of 29) in the 'within inflection' category: e.g. see 4a) and 4d). Lastly, we observed 2 instances of blending errors in the same category: e.g. 言われる /iwareru/ 'being said' and 呼ばれる /yobareru/ 'being called' were blended to form a non-word 言ばれる /ibareru/.

4) a. 基 つ く，b. 結婚式をあ た ました，c. 読みに か い，d. さけん た 時，e. 記述 ひ ながら

R-4 a. mo-to-<u>tu</u>-ku (mo-to-<u>du</u>-ku), b.ke-k-ko-n-si-ki.wo a-<u>ta</u>-ma-si-ta (a-<u>ge</u>-ma-si-ta), c. yo-mi-ni-<u>ka</u>-i (yo-mi-ni-<u>ku</u>-i), d. sa-ke-n-<u>ta</u>to-ki sa-ke-n-<u>da</u>to-ki, e. ki-ju-tu <u>hi</u>-na-ga-ra (<u>si</u>-na-ga-ra)

PM-4 a. /motozuku/'be based on', b. keQkoNsiki.wo/agemasita/ 'hold a wedding ceremony' c. /yominikui/ 'difficult to read', d. /sakeNda/toki 'when screamed', e. kijutu /sinagara/ 'while writing'

The next most noticeable errors were concerned with the presence or absence of /Q/ (geminated consonant), which is widely known as a typical L2 verb formation error, thus I strongly suspect that most of the samples in 5) occurred as a result of a grammatical formation error.

5) a. 出 っ て，b. 演奏し っ た，c. 言た，d. 可愛 がた，e. 通う て，f. 打ち勝 ち た

R-5 a. de-<u>t</u>-te (de-ϕ-te), b. e-n-so-u si-<u>t</u>-ta (si-ϕ-ta), c. i-ϕ-ta (i-<u>t</u>-ta) d. ka-wa-i ga-ϕ-ta (ga-<u>t</u>-ta), e. ka-yo-<u>u</u>-te (ka-yo-<u>t</u>-te), f. u-ti-ka-<u>ti</u>-ta (u-ti-ka-<u>t</u>-ta)

PM-5 a. /dete/ 'sticking out', b. eNsoo sita 'played', c. /iQta/ 'said', d. /kawaigaQta/ 'loved', e. /kayoQte/ 'commuting', f. /utikaQta/ 'overcame'

In summary, there were 25 geminate consonant related errors in the 'within inflection' category - 12 addition errors like examples in 5a) and 5b), 8 omission errors given in 5c) and 5d), and 5 substitution errors such as 5e) and 5f). A small number (n=8) of /N/ (moraic nasal) related errors such as 6) were also observed, majority (n=7) of which were found in the noun category.

6) a. めった ん に，b. 数年 か (間)，

R-6 a. me-t-ta-<u>n</u>-ni (me-t-ta-ϕ-ni), b. su-u-ne-<u>n</u> ka-ϕ (ka-n),

PM-6a. /meQtani/ 'rarely', b. suuneN kaN 'duringseveral years'

3.3.3 Semantic traceability by orthography types

Figure 3.2 displays semantic traceability of the intended words from erroneous words in each orthography type. While 81.2% (n=276) of errors were traceable, the remaining 18.8% (n=64) of errors were too heavily distorted to trace back the intended

meaning from a given error.

Figure 3.2: Intended word traceability by orthography
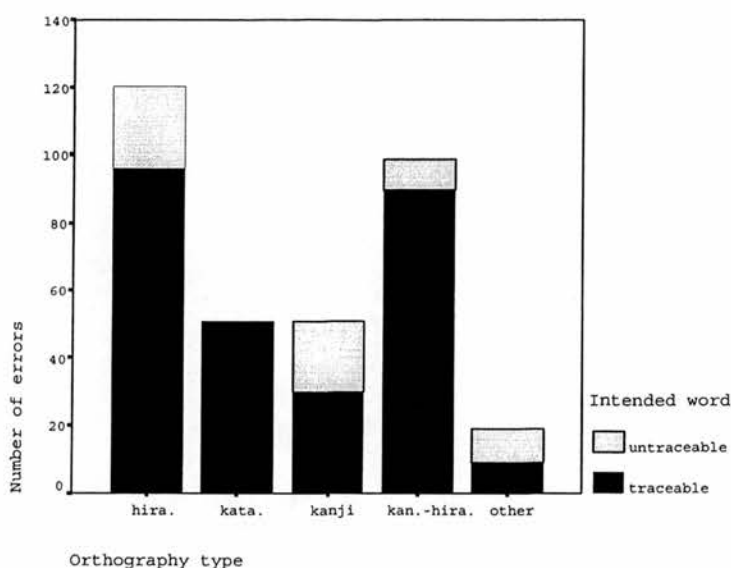
As shown in examples in 7), katakana errors (n=51) were 100% traceable.[7] Similarly, hiragana errors which consisted of a single error are traceable (n=96) shown in 8), whilst errors such as in 9) contain a hiragana chunk with multiple errors (n=24) that convey no meaningful units, hence making it difficult to identify the intended word(s).

7) a. イギ ル ス, b. ホムレス, c. バ ッ ブ, d. ハリ オ ッド, e. 彼の カ ラクター

R-7 a. i-gi-ru-su (i-gi-ri-su), b. ho-φ-mu-re-su (ho-—-mu-re-su), c. pa-b-bu (pa-bu), d. ha-ri-o-d-do (ha-ri-u-d-do) e. ka-re.no ka-ra-ku-ta-— (kya-ra-ku-ta-—)

RM-7 a. /igirisu/ 'Britain', b. /hoomuresu/ 'homeless', c. /pabu/ 'pub', d. /hariuQdo/ 'Hollywood', e. kare.no/kjarakutaa/ 'his character'

8) a. ちょ う と休む, b. 残虐 こい , c. むず が しい, d. びょ お き, e. や く そば

R-8 a. cho-u-to ya-su-mu (cho-t-to), b. za-n-gya-ku ko-φ-i (ko-u-i), c. mu-zu-ga-si-i (mu-zu-ka-si-i), d. byo- o -ki(byo- u -ki), e. ya-ku-so-ba (ya-ki-so-ba)

---

[7]We estimate agreement rate between the judgments of the author and of two native teachers of Japanese. They judged the same set, but each read the set in a different random order. The agreement rate was considerably high 83.8%. The agreement rate for each script type was as follows: hiragana (81.7%), katakana (94.1%), kanji (72.5%), kanji + okurigana (88.9%), and others (73.7%).

RM-8 a. /cjoQto/yasumu 'having a little rest', b. zaNgjaku/kooi/ 'cruelact', c./muzukasii/ 'diffi-
cult' d. /bjooki/ 'illness', e. /yakisoba/ 'fried noodle'

9) a. りこんしゃいた になる, b. 小さいなか 田舎の町, c. 怪獣に うったろう とした

R-9 a. ri-ko-n-sya-i-ta ni-na-ru, b. ti-i-sa-i-na-ka i-na-ka.no.ma-ti, c. ka-i-ju-u ni-u-t-ta-ro-u to.si-ta
RM-9 (All samples of the intended word in 9) are untraceable)

These examples suggest that we can generally identify the intended word from an
erroneous word with phonetic scripts hiragana and katakana apart from heavily
distorted errors illustrated in 9). Hiragana errors of type 8) occupy about a half
(49.2%) of the total hiragana errors in my error corpus. In addition to this, phono-
logically deviant kana errors are uncommon in L1 word-processed text. Thus, it is
probably reasonable to assume that kana errors given in 7) and 8) are very much
an L2 specific error. The analysis should then reflect the particular error features
among EJ writers.

First of all, we observed voiceless-voiced CV or CCV substitution errors such
as づる/zuru/ for つる /turu/ 'crane' and はんきゃく/haNkjaku/ for はんぎゃ
く/haNgjaku/ 'treason'. Secondly, vowel shortening such as /haNkoteki/ for /haN-
kooteki/ 'rebellious' and vowel lengthening examples like /uNteNsjuu/ for /uNteN-
sju/ 'driver'. This line of direct evidence should support my hypothesis concerning
the existence of EJ writers' spelling problems; in particular, EJ writers' phonologi-
cal input problem of the intended word. However, we may naturally ask ourselves
whether /uNteNsjuu/ appeared as the result of the writer's phonological difficulties
or lack of keyboard fluency, as the system requires a writer to hit the same key
twice to realize the second sequence of [u:]. Although it is often difficult to identify
the source of an error from the static written product, a persistent error pattern of
the same writer reveals that this was not a typing error - the error /uNteNsjuu/
was counted 9 times in total in the same composition, where we further observed
word medial vowel lengthening such as /kjoori/ for /kjori/ 'distance', as well as sub-
stitution errors between heavy syllables, for instance, /cjooto/ for /cjoQto/ 'just
a little'.[8] This example strongly suggests a fundamental cross-linguistic problem

---

[8]Strictly speaking, there is no relevant environment to replace the first member of geminate
/t/ with the second member of long vowel /oo/ in this example, but similar phenomena are known
among L1 studies, where a heavy syllable can be replaced with another heavy syllable without
any relevant environment (Kubozono, personal communication).

in perceiving and producing the L2 sound system. Further studies in the following chapters will address this issue of the relationship between L2 production and perception in Japanese with special reference to J-word processing.

While the vast majority of the above data seems to reflect writer's phonological input problems, the data also shows some instances (n=20) of semantically more transparent written errors that cannot be accounted for by wrong phonological input. For example, hiragana errors with regard to long vowels [oː], where there is not a one-to-one correspondence between sound and orthography, and thus a writer needs to choose the right hiragana letter between /u/ and /o/ according to the current hiragana usage. Furthermore, the long vowel [oː] attracted another L2 specific error such as せんとぅてき for せんとうてき 'aggressive' concerning the size of hiragana letter. This seems to be a case of mix-up with the written convention of geminate consonant /Q/, which is always marked by a small hiragana character. These orthographically motivated errors provide further evidence for another L2 specific writing problem, which was not reported in previous studies.

We have so far discussed semantic traceability of hiragana and katakana errors. Examples in 10) will illustrate written errors in kanji (n=51)

10) a. 代表 昨 (作), b. 復 雑 (複), c. 弟 一 (第), d. 間 単 (簡), e. 出 帆 (出版)

R-10 a. da-i-hyo-u- sa-ku , b. fu-ku -za-tu, c. da-i -i-ti, d. ka-n -ta-n, e. shu-p- pa-n

RM-10 a. /daihjoosaku/ 'masterpiece', b. /fukuzatu/ 'complicated', c. /daiiti/ 'the first', d. /kaNtaN/ 'easy',e. /sjuQpaN/ 'publication'

Of 30 instances of traceable kanji errors, 66.7% (n=20) were homophone kanji error as in 10) above. This evidence confirms a widely known L1-specific word-processed kanji error, which is normally portrayed as kanji-henkan miss or 'homophone conversion mistake', which is also found among EJ users of J-word processors. Remaining 10 other traceable errors (33.3%) include 5 instances of kanji error with a visual problem. I shall briefly touch upon untraceable kanji errors (n=21) presented in 11).

11) a. アフガニスタンの 前山 を旅行している, b. 祖谷 の老紳士, c. 卓話会

R-11 a. a-fu-ga-ni-su-ta-nn.no ma-e-ya-ma wo.rjo-ko-u-si-te-i-ru, b. so-ya.noro-u-si-n-si, c. ta-ku-wa-ka-i

RM-11 (All samples of the intended word in 11) are untraceable)

For instance, 11a) contains 2 very simple kanji 前 'front' and 山 'mountain'. We can understand the meaning of each kanji, but when they are combined as a 2-morpheme character kanji compound, it becomes a non-word with no corresponding sound representation in Japanese. As this example demonstrates, once a kanji like this appears in the text, it turns out to be more difficult to identify the intended word, because the phonologically opaque nature of kanji scripts means that a given kanji error normally cues much less phonological information of the intended word to the reader.

3.3.4   L2 kanji errors in digital text

Having presented the evidence for the existence of L2 word-processed kanji errors that are similar to the L1 problem of homophone kanji, let us now discuss the underlying phonological input in EJ kanji errors. As might be suspected from examples in 12), word-processed L2 kanji errors do not necessarily concern homophone kanji. Error kanji in 12) below revealed a certain degree of semantic anomaly, and they are undoubtedly not cases of pragmatic misuse. This is an important fact to stress - all kanji errors in 12) were based on partially deviant phonological input for the intended kanji.

12) a. 政|界中 (世界中), b. 愛|憎| (情), c. 諸|説 (小説), d. 師事| (詩人)

R-12 a. se-i-ka-i-ju-u (se-φ-ka-i-ju-u), b. a-i-zo-u (a-i-jo-u), c. sho-φ-se-tu (sho-u-se-tu), d. si-ji-φ (si-ji-n)

RM-12 a. /sekaijuu/ 'all over the world' b./aijoo/ 'love' c./sjoosetu/ 'novel' d. /sijiN/ 'poet'

A first glance at these kanji in 12) gives an impression that they are of homophone kanji errors similar to examples in 10), whereas Roman transcripts of each error with the correct Roman inputs in R-12 reveal that the errors in 12) did not share the identical underlying sound with the intended word. Instead, there is a common feature among the errors in 12) that they included only a fraction of phonological mistake that was very close to the right version. For instance, the error 政界中 'throughout the political world' in 12a) was made by adding a short vowel /i/ that formed a sequence of a long vowel [e:] as se-i-ka-i-ju-u. As a result of this insertion, a list of at least 6 homophone kanji must have been displayed on the screen. Then, 政界 'political world' was probably chosen in place of 世界 'the world' as both compounds feature the common second element 界 'society/boundary'. This analysis accounts for the difference between a single visual mistake of homophonic

kanji error and a multiple phonological-visual mistake of kanji errors such as in 12). An analysis of 'Reading type' yielded a distribution of non-homophonic errors within hiragana and kanji categories as 64.3% (36 out of 56) in on-reading and 45.6% (31out of 68) in kun-reading. The bulk of non-homophonic errors in both reading types strongly indicates the reality of an L2-specific spelling problem. In addition to this, as assumed by the designers of J-word processors, L1 users normally do not make phonological input errors. All these things make it clear that kanji errors illustrated by 12) can be identified as L2-specific word-processed kanji error.

As for kanji errors alone, Table 3.4 gives number and percentage of kanji errors according to whether errors were homophonous.

Table 3.4: Number and Percentage of Kanji Errors (n=51) by Category 4

Homophones	20	(39.2%)
Non-homophones	14	(27.5%)
Unclassifiable	17	(33.3%)

As has been discussed, unclassifiable kanji errors (n=17) illustrated by 11) included no clue to make a judgment if a given error was homophonous. However, kanji errors judged as homophones had 100% semantic traceability, so that the chance of adding more homophone examples from unclassifiable kanji error seems rather low. In summary, homophone kanji errors like those given in 10) take up 39.% (n=20) of the total kanji errors, while the total number of non-homophonic kanji and unclassifiable errors fill a larger portion of 60.8% (n=31) in my kanji data.

The examples presented in 13) and 14) will provide another line of evidence that wrong phonological input was not the only cause for the appearance of non-homophonic kanji errors, but other factors were also involved. In addition, multiple factors may have played a role in the errors given in 13) and 14). For example, 13a) involved a mixture of 3 errors: first, partially deviant phonological input occurred, secondly a word segmentation problem occurred, and then a visual recognition error was made.

13) a. 結婚を 洩仕 込む (申し), b. 英田 (間), c. 受け 手 いる (て), d. 蟹座 (癌)

R-13 a. ke-k-ko-n.wo mo-ra-si-ko-mu (mo-u-si-ko-mu), b. e-i-da (a-i-da), c. u-ke-te i-ru (u-ke-te) d. ka-ni-za (ga-n)

RM-13 a. keQkoN.wo /moosikomu/ 'propose a mariage', b. /aida/, 'during' c. /ukete/iru, 'receiving' d. /gaN/ 'cancer'

Similarly, while underlying phonology of 13c) is identical to a gerund form of the intended verb u-ke-te-i-ru, a conversion must have been made at the wrong segmentation point at u-ke-te that resulted in the appearance of a noun 受け手 'receiver' followed by a verb いる i-ru 'be'.

The example in 13d) involves no common phonological inputs between the error /kaniza/ and the intended kanji /gaN/. It is an interesting case as the examples we have discussed so far have been cases of partial phonological anomaly in error words. We were fortunate that the female writer of this error kanji was available for 'authoritative interpretation' (Corder, 1973, p.274). According to her explanation, she wanted to write 'James Dean's mother died of cancer' in Japanese, but she did not know the Japanese word 'cancer'. She then used an on-line bilingual dictionary, which gave all the Japanese translations of 'cancer' including the serious disease cancer 癌 /gaN/ and an astrological star-sign the Crab 蟹座 /kaniza/. Regrettably, she picked up the latter being convinced herself that the on-line dictionary provided her with synonyms of 'cancer'; hence, no attempt was made to check the suitability of the chosen word by reversing it into English, and she made a cut-and-paste directly from the on-line dictionary. A more advanced writer could have chosen the right kanji by identifying a semantic element to cue 'illness' in the correct kanji. It was after all a translation mistake as a result of the writer's linguistic weakness on kanji discrimination ability. In any case, without the writer's explanation, it was impossible to arrive at what she wanted to convey by writing 'James Dean's mother died at the Crab'. In short, this particular error clearly demonstrates that a word-processed kanji error can be written without typing the sound of the intended word.

14) a. 息 話 まる (詰)，b. 持 に (特)，c. 下 真面目 (不)，d. 彫刻 版 (板/ 版)

R-14 a. i-ki-wa/hanasu-ma-ru (i-ki-zu-ma-ru), b. ji/motu.ni (to-ku.ni), c. ge/sita-ma-ji-me (fu-ma-ji-me) d. cho-u-ko-ku-ha-n (cho-u-ko-ku-ba-n)

RM-14 a. /ikizumaru/ 'breathtaking', b. /toku.ni/ 'specially',
c. /hu.majime/ 'not serious', d. The intended version cannot be retrieved

Looking at the errors placed in 14) remind us about a typical hand-written error with visual representation problem: i.e. visually very close, but sound and meaning are very distant from the target. These errors generally involving some small component of the whole kanji configuration is wrong by one or two strokes, which create an almost identical visual representation of kanji or an invented kanji to the intended target. The writer can write such kanji errors without knowing pronunciation of kanji by hand. On the other hand, as we can see from R-14, the writer

needed to type the sound of the error words. Thus, it is apparent that the errors in 14) do not share the same characteristic with the similar hand-written errors. For instance, the configuration of a non-word 14a) 息話まる is visually very similar to the intended kanji 息詰まる 'breathtaking', yet there is no common underlying phonological input between the error constituent and the target portion. In order to write the erroneous version, the writer must intentionally enter either ha-na-si (ku-reading) or wa (on-reading) to obtain 話 'story' after converting the first element i-ki (息) into kanji instead of keying in i-ki-du-ma-ru for the right version. It is obvious that writing the erroneous kanji requires more time-consuming keyboard operation than for the right version. While we can suspect the involvement of visual factors in these examples, Error Analysis cannot account for why and how EJ writers wrote such awkward errors in 14) by deliberately decomposing the intended kanji compound. Since no comparable observation has been reported in the previous literature, the following chapters will pursue this type of visually similar word-processed kanji error along with the issue of EJ writer's phonological input problems.

3.4 Summary

In this chapter, I have analyzed 340 orthographic errors by EJ writers with main emphasis on spelling problems in L2 J-word processing. Various lines of evidence strongly suggest that EJ writers faced particular problems at the initial phonological input stage of J-word processing, thus confirmed that EJ writers were not liberated from lower-level routine spelling tasks when writing with J-word processors. The error pattern in Category 1 showed the high proportion of hiragana errors indicating that there were certain L2-specific spelling errors. Firstly, errors with a partial deviation of phonological input revealed several key factors. For example, vowel deletion and addition errors indicated a cross-linguistic problem regarding the durational contrast of Japanese vowels. Secondly, many voiced-voiceless CV substitution errors pointed to the involvement of phonological phenomena of sequential voicing, especially an error prone item was an orthographic distinction between alternative hiragana letters ず (zu) and づ (du) that share the identical underlying sound [zu]. Similarly, we observed an influence of Japanese orthography in errors that were based on confusion between two-to-one hiragana-sound correspondences: e.g. う/u/ or お/o/ for the second sequence of [o:]. Finally, we have observed some errors concerning semantic blending and visual recognition problem. All in

all, 4 factors above played a crucial role for the appearance of a high proportion of hiragana errors in my corpus.

The second type of collective feature was found in kanji errors. We observed that homophone kanji error formed one category in EJ kanji error, hence confirmed that well-known L1 phenomenon of homophone kanji error was also held in L2 corpus. However, a difference was that there were more non-homophonic errors than homophonic kanji errors in my EJ corpus. Accordingly, we have found evidence of L2 specific kanji error pattern in non-homophonic kanji errors involving multiple errors based on phonological and visual factors. A typical EJ error was a small mistake at the initial phonological input, which was followed by a detection failure of the wrong kanji. It is normally difficult to infer the intended kanji from this type of error, because the error kanji not only indicates a completely different meaning, but also forces the reader to recover the whole word phonology of the intended kanji from partially available phonological information in the error kanji. Moreover, there were translation mistakes that were caused by the writers' lack of kanji discrimination ability when use on-line look-up functions. Finally, we found some non-words that were similar to hand-written errors on the surface. Even so, our closer examination revealed that these visually similar word-processed kanji errors were not written in the same way as similar hand-written errors are normally written. Although it is interesting to find out the reason why L2 writers intentionally wrote such kanji as if they were written by hand, this issue has never been addressed by previous studies since almost every study concentrated on phonological aspects of L2 J-word processing errors. The next chapter will deal with kanji errors of this type, especially focusing on why EJ writers find it necessary to write each constituent of a kanji compound separately rather than writing whole word units at word level.

CHAPTER 4

FACTORS IN L2 KEYBOARD SPELLING ERRORS

This chapter attempts to determine how L2 writers write with J-word processors, and why and when word-processed orthographic errors occur. Traditional Error Analysis is not very helpful to answer these questions because it deals with the end product. As Ellis (1985, pp.51-54) points out, Error Analysis can provide only limited information when we search for the prime cause of errors. In addition, while careless slips (i.e. typing errors) must be taken into account as a factor of machine-mediated writing errors (Rumelhart & Norman, 1982), we cannot trace the writing process of any given error by Error Analysis. Therefore, the current chapter aims to identify more specific factors through direct observation of writers' on-going writing processes, and their cognitive activity while they are writing with J-word processor. For this purpose, recorded samples of concurrent protocol data (Ericsson & Simon, 1979) were taken from the vocalized thoughts of L2 writers (n=5) at various levels, and from native Japanese (L1) participants (n=2) while they solved the tasks. The transcribed protocol data were analyzed to determine L2 users' strategies when writing with J-word processors. In addition, strategic differences between L1 and L2 users in their kanji error monitoring tasks were brought into light. It should be stressed that the experiments presented in this chapter should be seen as pilot studies. Although they used a small number of participants, and did not employ statistical analysis, we will see that the pilot data can be useful for generating new research questions that can be addressed experimentally.

4.1 Keyboard skills

4.1.1 Method

The first study (Experiment 1: hiragana copy task followed by kana-kanji conversion task) aims to identify causes for the appearance of orthographically illegal hiragana strings that are left unconverted into kanji. Tasks are designed to observe how L2 writers write given hiragana words in a phrase or a short sentence, from which I expect to uncover both random and systematic error sources.

Participants

Two female and one male native speaker of English took part in the experiment for a small payment. They participated in the study after taking 225 hours of beginning classes including 18 one-hour computer-writing classes at the University of Edinburgh.

Materials

10 phrases and 20 short sentences (number of words, including nouns, verbs, adjectives and post positional particles, per phrase or sentence are M=3.2, SD=0.63; M=3.1, SD=1.07 respectively) were created using selected words that involve previously identified writing problems. The first type of writing problem is concerned with the distinction between voiced and voiceless consonants. The second type of writing problem involves heavy syllables - namely those that contain geminate consonants (transcribed as /Q/), moraic nasals (transcribed as /N/), and the second vowel in an identical vowel sequence (transcribed as /VV/), all of which are represented by one hiragana letter. The third type of writing problem involves a one-mora-two-hiragana type sequence (transcribed as /CjV/) such as きゃ /kja/. 6 sentences are 'straight' containing neither heavy syllables nor voiced consonants except the subject marker ga, while remaining 24 sentences contain at least one of the above problematic items.

4.1.2　Procedure

The participants were asked to re-type the given hiragana phrases or sentences, and then convert them into hiragana-kanji mixed text (see Appendix D). A PC with word processing software Mocrosoft Word 2000 [1] was used to run the experiment. Lotus ScreenCam97 was used to record real-time screen movements as a dynamic movie. A Coomber 393 Recorder and PZM microphone were used to record concurrent verbal protocol of each participant. Recorded protocol data were transcribed by the experimenter with SANYO MEMO-SCRIBER (TRC-8080).

The participant's writing processes were observed in three ways: (1) tape recorded think-aloud protocols, (2) real-time screen movements, and (3) the experimenter's field notes that were taken during each session. Three participants were tested individually or in pairs as a writer and an advisor. It was originally intended to form two groups with four participants, but we had to carry out the experiment with three participants instead, due to a cancellation at short notice. The experimenter instructed participants to verbalize their guesses, hypotheses and the knowledge they were using while completing their writing tasks, and the pair was instructed to discuss the matter whenever a problem arose during the task. There was no time limit for completing this task in order to reduce the risk of participants being nervous, which often results in a much lower occurrence of think-aloud protocols (Kaiho and Harada 1993). The experimenter made it clear to participants that she was more interested in observing their writing processes than the final products, so that it was important to verbalize their thoughts as often as possible.

4.1.3　Results and discussion

Repair units, L2 keyboard skills and their linguistic knowledge

The data was analyzed by measuring the number of repair units. My operational definition of a 'repair' is one hiragana input error that is followed by a delete and a rewrite. The analysis yielded that the ratio between error and error-free sentences or phrases was 2:28 in both groups, and they shared an identical error for item No. 9 (see Appendix D): the one mora /CjV/ sequence in びょういん /bjo.o.i.N/ (hospital) was written instead of the target two morae /Ci/-/jV/ sequence びよ /bi.jo/ of び よういん /bi.jo.o.i.N/ (beauty salon). A further protocol analysis revealed that

---

[1]The 'Auto correct' function of MS-IME was disabled during the experiments in this chapter.

both groups read the target /bi.jo.o.i.N/ (beauty salon) as /bjo.o.i.N/ (hospital) immediately they saw the word, and their concurrent protocol 'That's hospital' confirms that it was intentional rather than a slip or a typing mistake. They then converted the hiragana string into 病院 (hospital) without attempting to make any repair. This implies a strong case of a visual discrimination problem or an influence of the degree of semantic relatedness: the wrong reading 'dog hospital' may be perceived as more plausible than the target 'dog beauty salon', and this may have led participants to read the given word differently.[2]

Similarly, in item No.16 (see Appendix D), there was an instance of a mix-up between two near-synonymous suffixes: the agentive suffix -しゅ(手: hand) /sju/ and -しゃ(者: person) /sja/ in うんてんしゅ(driver) /u.N.te.N.sju/. We can see from these instances that some beginning writers' errors may consist of a copy mistake due to semantic influences. Clearly, there was no phonologically based error source in the above written errors. This finding may add a new written error source to those of Kondo's durational study. According to Kondo (1998, p.31), a /CjV/-/CijV/ altering written error is a typical mistake even among advanced learners' writing, whose language background (native speakers of British English) is similar to those in the current study. She reported that both beginners and fluent learners of Japanese made two types of pronunciation errors involving /CjV/ sequences - (1) inserting /i/ between /C/ and /j/, or (2) deleting /i/ from a /CijV/ sequence. Nevertheless, she seems to suggest that /CjV/-/CijV/ altering written errors occur owing to the single factor of phonological influence. While this may be the case for most of the written errors of this type, the current study points out that we cannot ignore other factors such as semantic influences in the written error production.

There were 36 (60%) repair-free instances, all of which were also error-free except for the items described above. So far, we have seen that the total of 34 (pair : solo=15:19) correct versions without any repair take up 56.7% in this task. Figure 4.1 shows the percentage of trials in which the solo and pair participants made zero repairs, one repair, etc, up to 12 repairs. The highest point of both writers is marked by repair-zero category, and then there is a rapid drop at repair-once category in both pair and solo trials. This indicates that L2 novice writers' keyboard skills were reasonably competent, and we are likely to find frequent typing problems within the repair-once category. Positively, a total of 5 instances of slips were found in repair-once (4 times) and repair-twice (1 time) categories. Although participants had no problem in producing the geminate consonant itself, there were 3 counts

---

[2]Dog's hospital (a vet) seems more common than dog's beauty salon in Britain.

Figure 4.1: Number of repairs by L2 novice writers
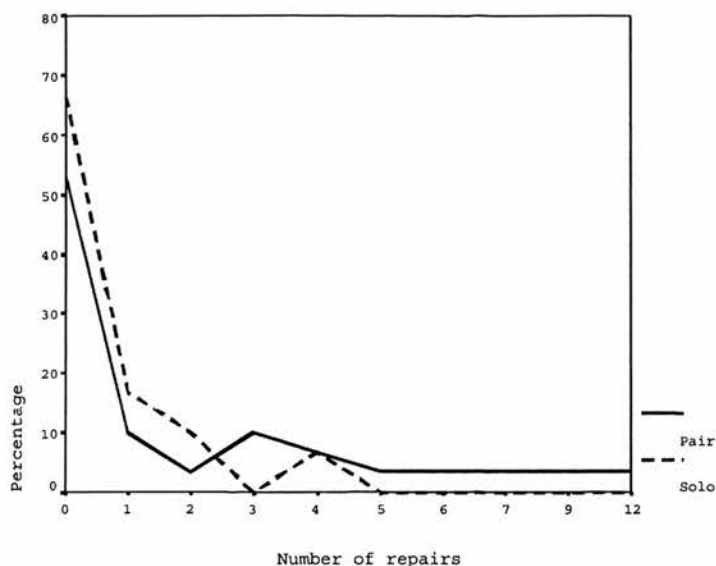
of slips for the following character: /kek.ke.N/ for /kek.ko.N/, /ik.ka.ga.tu/ for /ik.ka.ge.tu/, /tta/ for /tot.ta/. This seems to indicate that the segment(s) following a geminate consonant is prone to a slip. One explanation for this observation may be that producing a small hiragana letter っ as the realization of the geminate consonant requires the writer to hit the same key twice, which may trigger a slip in the following item. However, we may safely eliminate the possibility of a simple keyboard error for the geminate consonant itself in a further production study with the same participants.

Identifying the systematic sources for digital misspellings

The point to observe is that most of the slips were found in the repair-once category, so that if we closely investigate twice or more repaired items, we may be likely to identify more problematic and systematic sources for word-processed written errors. We shall now look more carefully into the number of repairs on each item. A total of 88 (86.3%) targets were error-free, while 13 (12.7%) targets attracted 1 to 12 repairs.

From the table 4.1, three writing error sources: (1) slips concerning a geminate consonant, (2) orthographically illegal hiragana, and (3) mechanical typing error involving moraic nasal, were found across paired and solo participants. First, the

Table 4.1: Repaired items (two or more repaires)

Target word	Error item	Number of repairs			Error source
		pair:solo	M	SD	
まんいん (ma-N-i-N :'jam-packed')	まにん (ma-<u>ni</u>-N)	9:4	6.5	3.54	assimilation
にほんの (ni-ho-N-no: 'of Japan')	にほん<u>お</u>(ni-ho-N-<u>o</u>)	7:2	4.5	3.54	assimilation
おんな (o-N-na: 'woman')	おん<u>あ</u>(o-N-a)	3:1	2	1.41	assimilation
みかづき (mi-ka-du-ki: 'new moon')	みか<u>ず</u>き (mi-ka-<u>zu</u>-ki)	2:4	3	1.41	orthography
いっかげつ (i-Q-ka-ge-tu: 'a month')	いっか<u>が</u>つ (i-Q-ka-<u>ga</u>-tu)	12:2	7	7.07	multiple
ちょうしょく (cho-R-sho-ku: 'breakfast')	<u>しょ</u>うしょく (<u>sho</u>-R-sho-ku)	3:0	-	-	slip
おとうさん (o-to-u-sa-N: 'father')	おと<u>お</u>さん (o-to-o-sa-N)	4:0	-	-	orthography
とった (to-Q-ta: 'took')	とた (to-ta)	1:0	-	-	slip

participants had a problem in writing いっかげつ/ik.ka.ge.tsu/. The solo participant omitted the geminate consonant as いか/i.ka/ at her first attempt, but she immediately noticed the slip and made a repair, but the second slip was made in the following item. In the case of the paired participants, they successfully wrote the geminate consonant without making any repair, but they made a mistake in the following CV sequence between two possible readings in 月 /ge.tsu/ and /ga.tsu/ (month), which led to 12 repairs until they realized the reason why they could not convert the hiragana string いっかがつ /ik.ka.ga.tsu/ into kanji.

Secondly, an illegal hiragana problem was observed in a sequential voicing item づき (du.ki) of みかづき/mika.zuki/. There were four items むかしばなし (old tale) /mukashi.banashi/, ゆきぐに (snow country) /yuki.guni/, ひとびと (people) /hito.bito/, and みかづき (new moon) /mika.zuki/ that tested how the participants write sequential voicing targets. The first three targets were written straight away without difficulty, whereas the last item attracted 2 repairs by the paired participants and 4 repairs by the solo participant. The reason was obvious - there is only one hiragana to represent the underlined voiced CV in the first three targets (see above), while two different hiragana ず (zu) and づ (du) are available for the fourth target, and a writer needs a morphologically based spelling strategy in order to choose the right version づ (du), which is related to つ (tu) morphologically. Recorded screen movements revealed that both participants first wrote the default-version ず (zu). The writer in the pair stopped when she wrote as みかず mi-ka-zu, and asked her partner if he knew how to type the voiced version of つ (tu) that is the underlined sound of the initial consonant of the target. The following episode tells us that they were aware of the internal structure of the word, but they did not know exactly how to type the target hiragana.

1. W: How can you type /zu/ (づ)?
2. A: I can never do it.
3. A: Right, if you press 'd' and 'u'⋯ [づ appears on the screen]
4. W: Oh! [laugh]
5. W: It suppose to be like ず (zu).
6. A: I'll try to use the 'd-u'. [typing as 'mi-ka-du-ki' and pressed the kanji conversion key]
7. W: It just gave us 三日月 (new moon).   (W=writer, A=advisor)

The solo participant was also conscious about the two different hiragana versions ず (zu) and づ (du). She said that she had no idea how to type づき/zu.ki/, but she remembered that 月 /tu.ki/ was sometimes read as /zu.ki/. Accordingly, she divided the word into two segments and typed 'mik-ka' for 三日 (three days) and 'tu-ki' for 月 (moon) to get the compound kanji 三日月 (new moon). This analysis confirmed that our participants' problem was their typing skill rather than linguistic knowledge on sequential voicing or lack of morphological awareness.

Similarly, there was a mistake concerning orthographic choice that shares an identical phonetic value. According to my analysis on each vowel target, L2 writers wrote long vowels /aa/, /ii/, /uu/, /ee/ straight away without making any mistake. The only item that attracted several repairs was /oo/ in おとうさん /o.to.o.sa.N/. The writer first omitted the target vowel, then in the following repair episode orthographically illegal お /o/ appeared in the place of required target う as in とう/to.o/. This seems to me a reasonable result, because long vowels /aa/, /ii/, /uu/ have one and only one (repeated) hiragana letter, so that the writer can always be sure about the target vowel. In the case of long vowels /ee/ and /oo/, there is an alternative hiragana letter that shares the identical phonetic value. The words including the former type vowels are regular words, whereas the words with the latter type vowels are ambiguous orthographic representations. However, discriminating two versions of [e:] 'e-e' and 'e-i' is fairly easy, because the number of lexical items that contain the sequence 'e-e' is extremely few, and L2 writers can almost always safely write the 'e-i' version. On the contrary, there are more items that consist of 'o-o' or 'o-u' for the long vowel [o:]. Thus, this particular item seems worth investigating further in following studies.

Thirdly, the participants had a problem concerning the moraic nasal /N/ within a word as well as across a word boundary. I have identified the underlying cause as a typing problem similar to the difficulty in obtaining the geminate consonant

described above. First, I eliminated words containing /N/+CV sequences such as り
んご (apple) /ri.N.go/, since these items were not prone to an assimilation problem.
In contrast, /N/+/nV/ or CV-/N/-V sequences automatically invite an assimilation
with the following non-moraic nasal or vowel, if the 'n' key is hit only once. For
example, the CV-CV-N hiragana string まにん /ma.ni.N/ is found instead of the
correct まんいん (jam-packed) /ma.N.i.N/. According to my protocol analyses,
both writers pronounced the target word properly: i.e. まんいん was pronounced
as /ma.N.i.N/ rather than /ma.ni.N/ while they type in the target word. Yet,
their typing resulted in an unwanted CV-CV-N hiragana string まにん /ma.ni.N/,
so that one of the participants commented on this as 'It's gone crazy'. Although
they were taught to hit the 'n' key twice for ん /N/ at the first computer writing
session, the writers did not seem to remember the given instruction. This finding
may suggest that similar L2 word-processed writing errors are likely to occur if the
writer is unaware of this mechanical requirement for pressing the same key twice to
generate a geminate consonant or a moraic nasal.

4.1.4 Summary of preliminary data

We can sum up what we see as emerging from the above analysis as follows:

1. L2 writers had no major problems in writing either voiceless and voiced CV
   or CjV type hiragana, if the target letter has a one-to-one grapheme-phoneme
   correspondence, but beginning writers may have a visual discrimination (copy)
   problem between visually similar hiragana letters.

2. L2 writers can type the geminate consonant sequence correctly, but this item
   seems likely to trigger a slip in the following segment due to a typing require-
   ment to hit the same key twice for generating the geminate consonant.

3. L2 writers experienced difficulty in typing a voiced CV type hiragana with
   two-to-one grapheme-phoneme correspondence.[3]

4. L2 writers wrote long vowels /aa/, /ii/, /uu/, /ee/ straight away without
   making any mistakes. The only item that attracted a repair was long vowel

---

[3]Fortunately for them, there are only two sets of voiced CV hiragana (ず/zu/ , づ /du/ and
じ /ji/, ぢ/di/) that share the same phonetic value, so that learning how to type differently is
relatively easy for them.

/oo/.

5. A typing problem with moraic nasal was found across participants. Because they hit the 'n' key only once instead of twice, the sequence /N/+/nV/ or CV-/N/-V mechanically invites an assimilation. Although this is a simple typing error, L2 writers suffer from this assimilation phenomenon a great deal, because it leads to a new segmentation problem if assimilation occurs across the word boundary.

All in all, we could reduce the errors into 3 categories. To begin with, a copy problem that concerns the writer's visual discrimination ability. There seems to be a contextual effect in such cases. Next, a typing skill that requires a writer to hit the same key twice in writing one hiragana letter for moraic nasal or geminate consonant. Lastly, orthographic knowledge about alternative hiragana letters that share the same phonetic value. Except in the first case, we could easily incorporate the second and the third findings into our computer-based writing instructions.

4.2 Retrieval memory and knowledge of orthographic conventions

Experiment 1 (hiragana copy task in 4.1) focused on participants' typing problems when L2 writers retype given meaningful hiragana chunks. However, we should note that L2 writers do not necessarily need to search through their mental lexicon to recall the sound of the target words in Experiment 1 as phonetic script hiragana clearly cued the word sound. Experiment 2 (hiragana production task; see Appendix E) is designed to address exactly this problem. Tasks were designed to include no phonological information concerning the target words. Subjects were asked to recall the whole word phonology of a target word via meaning (i.e. given English word) and surrounding context that was written in kanji-kana mixed Japanese. I expected to be able to make a direct observation of L2 production problems, which are likely to result in the appearance of orthographically illegal hiragana strings unconverted into kanji.

4.2.1 Method

Participants

The participants of Experiment 2 were the same three English native learners of Japanese who participated in Experiment 1.

Materials

First, I selected 31 target Japanese words that include heavy (bimoraic) syllables that contained /N/ (n=4), /Q/ (n=4) or /VV/ (n=11)[4]. Some light (monomoraic) syllable words (n=14) were also chosen as fillers (n=9) or to form minimal pairs (n=5), which allow us to test whether participants have any problem in distinguishing a given pair in their recall and production. The following minimal pairs are concerned with two potential writing problems involving geminate consonants or long vowels were selected: (1) /ki.te/-/kit.te/, and /ki.ta/-/kit.ta/ (geminate consonants), and (2) /i.e/-/i.i.e/, /ki.te/-/ki.i.te/, /to.ri/-/to.o.ri/, /is.sjo/-/is.sjo.o/ (long vowels). An orthographic problem, which was identified in the previous section, regarding the writer's ability to distinguish hiragana letters with the same phonetic value /oo/ in writing will be investigated. Selected items were とおり- (*to pass*) /to.o.ri/ and こおり (*ice*) /ko.o.ri/. Next, the target Japanese words were translated into English. The main reason for using English words is that phonetic hiragana letters are not suitable as stimuli in this recall task, whereas using English words can mask the whole word phonology of the target Japanese words, simultaneously providing subjects with a semantic route to reach the target word. Translated target words were then embedded into 30 kanji-kana mixed sentences. On the whole, there were 10 short sentences that consist of 2 to 3 morae CV or CjV words. The remaining short sentences including at least one of heavy syllables /N/, /Q/ or /VV/.

4.2.2 Procedure

All procedures for Experiment 2 were exactly the same as for Experiment 1, except that the task was to read the given sentences once, then replace English word in

---

[4]/is.sjo.o/ (one's life) and /is.sju.u/ (one week) includes /Q/ and /VV/.

brackets with word(s) in kanji according to the instruction given (see Appendix E).

4.2.3 Results and discussion

Both solo and paired participants answered 80.6% correctly. They did not experience any problem with moraic nasal targets regardless of position[5]. Next, both writers made a slip with a geminate consonant, but managed to reach the target without making further repairs. The paired participants answered every task, but made an error on a geminate consonant for item No. 8 (see Appendix E). This error was made such a way that the stem of the non-past form /ki.ru/ (*to cut*) was first written, then inflectional ending /ru/ was deleted, resulting in the kanji stem 切. The writer then typed the past inflectional ending /ta/, which caused the appearance of both ungrammatical and orthographically illegal non-geminated versions 切た /ki.ta/ for 切った /kit.ta/ (*cut-past*). Interestingly, the same writer wrote another minimal pair, /ki.te/ (come-gerund) - /kit.te/(stamp) perfectly. If the writer had typed the target as 'kitta' (*cut-past*) like she did it for non-derivational 'kitte' (*stamp*) for item No. 24, the machine would have given her the desired target immediately. Previous researchers (Komori, 1998; Sakamoto, 1993) on L2 Japanese verb formation errors point out that the omission or addition of a geminate consonant occupies a large portion in te-form (gerund) errors. For instance, Komori (1998) reported that 64.3% of te-form errors in her study were either of omission or addition of a geminate consonant. In my L2 error corpus, the same category takes up 89.5% of the verb error category though I could not pin down the root cause for this error type from various possible error sources such as a morpheme alternation mistake, a typing mistake (slip) or L1 phonological interference. Together with the finding in the previous section, I can now consider this particular case to be a grammatical formation error. Although it is risky to draw a general conclusion from an instance, it seems likely that a combination of grammatical formation problems and slips increased the number of errors of this category in my L2 corpus.

The following table 4.2 is a list of problematic targets that attracted 1 to 2 repairs before kanji conversion took place.

---

[5]For Experiment 2, assimilation-inducing items were not used.

Table 4.2: List of targets with 1 to 2 repairs

Target (English)	Target	Error	Repair
cut-past form	きった ki-Q-ta	切た ki-ta	きった ki-Q-ta
pass	とおり to-o-ri	とうり to-u-ri	—
Tokyo	とうきょう to-u-kyo-u	とうきょ to-u-kyo	とうきょう to-u-kyo-u
one week	いっしゅう i-Q-shu-u	いっしゅ i-Q-shu	いっしゅう i-Q-shu-u
singer	かしゅ ka-shu	かっしゅ ka-Q-shu	かしゅ ka-shu
		かしゃ ka-sha	—

Semantic and visual confusion and sub-lexical kanji writing strategies

I have so far discussed a problem regarding geminate consonants. Next, I shall briefly outline a semantic error concerning one of filler items /ka.sju/ ('*singer*'), because both writers made 1 to 2 repairs on this item. It was a slip for the solo writer, while the writer of the pair intentionally typed it as 'kasha', from which she could not get the target kanji 歌手 /ka.sju/ ('*singer*').

She instantly broke down the kanji compound into two constituents, then typed the kun-reading of each kanji 'uta' (*sing*) and 'mono' (*person*), so that the converted single kanji 歌 and 者 respectively will form the intended kanji compound 歌者. Now, we need to recall a similar item in Experiment 1 - there was an instance of a mix-up of near-synonyms, involving the agentive suffixes - しゅ(手:*hand*) /sju/ and - しゃ(者:*person*) /sja/ in うんてんしゅ (*driver*)/u.N.te.N.sju/. The solo writer made this error. According to Shibatani (1990, p.219), Sino-Japanese agentive suffixes like -shu (手) and -sha (者) have no obvious distribution pattern, so that we need to remember them one by one. Both of our participants showed the developmental nature of their vocabulary learning in this aspect. So far as the current study is concerned, I have found enough evidence to claim that L2 novice writers do not always write kanji compounds in the same way as that of L1 users of J-word processors. L1 writers normally achieve the intended words by typing the whole word phonology of the target kanji, thus, seldom break down a kanji compound into two constituents. This is the most natural and effective way to write with J-word processors. Although L2 writers do write in a similar manner, my observation has so far revealed that they occasionally decompose a kanji compound on purpose, and then try to write individual characters to form what they believe to be their target kanji compound.

Failure to notice visually similar kanji

Turning now to long vowels. The writers did not make repairs on long vowels /ii/, /ei/ and /ee/, whilst each of the geminating vowels /uu/ and /oo/ invited 2 occurrences of error-repair episodes. As we can see from the table 4.2 above, one item concerns the two-to-one correspondence of long vowel /oo/. The solo writer typed to-u-ri for to-o-ri, and in so doing she could not convert the hiragana string into kanji, then she abandoned this task by saying that she did not know the word. There is one more task ko-o-ri (*ice*) that requires 'o-o' version, but both participants could not recall the pronunciation of this word. The solo writer went for *ice* in katakana as a loan word アイス '*aisu*'. The participants in the pair also could not recall the target pronunciation, so they discussed the shape of this kanji. They said that the shape of the kanji *ice* (氷) looked very similar to that of *water* (水).

The participants in the pair then decided to search for the target *ice* (氷) without taking a phonological pass, and they asked the experimenter if they were allowed to use the IME pad so that they could search the target by counting its total stroke number. IME pad provides a handwriting recognition pad, on which the writer can use the mouse to write desired kanji as if writing with a notebook and pen. This function is useful when the writer knows the shape of the kanji but cannot remember its pronunciation. It was an interesting case, because the novice writers knew an alternative way to get the desired kanji when they know the shape of the kanji without knowing how to read it.[6] Using this function, they managed to find a list of kanji that includes both water (水) and ice (氷), but they failed to recognize the desired kanji among 14 other kanji, all of which include water (水) as a component. In spite of everything, their kanji recognition ability was not sufficient for this type of search. In short, their protocol data tells us that they apparently had the form-meaning connection, with which they could have written the target kanji by hand, yet a combination of lack of phonological information and their insufficient visual recognition skill made this task impossible to write with the J-word processor. It offers the key to an understanding of L2-specific recall problems with J-word processors, and it is worthwhile examining this crucial issue more closely in a later study in Chapter 7.

---

[6]Finding the desired kanji by IME pad takes much longer time than the routine method, but it is a useful alternative way to get the desired kanji even for Japanese writers.

Vowel deletion errors and potential underlying error sources

I shall next discuss an omission of the second vowel in the vowel sequence in word-final position in the case of /to.o.kjo/ for /to.o.kjo.o/ and a similar case of omission in /is.sju.u/ (*one week*). In Experiment 1 (copy task), L2 writers made no copying mistakes with long vowels /aa/, /ii/, /uu/, /ee/, in contrast to /oo/, which attracted several repairs. In the current production study, the writers again made no mistakes on /ii/ and /ee/ including its orthographic variation 'e-i'. From the observations above, we can be fairly certain that L2 novice writers can handle long vowels /ii/ and /ee/ rather well. In addition, L2 writers could be more confident with words involving long vowels /aa/ and /ii/. The reason is that Japanese origin words including the geminating /aa/ are limited to very few family terms such as 'okaasan' (*mother*), and 'obaasan' (*grandmother*). Moreover, the need of hiragana い /i/ as the second vowel in the vowel sequence /ii/ is often obvious to the L2 writers when they write so-called i-adjectives in Japanese, which always end with い/i/.[7] For example, ながい /na.ga.i./ (*long*) and おいしい /o.i.si.i/ (*delicious*) are adjectives of this type. In the case of the latter, we can find an instance of the geminating vowel sequence /ii/, and it is likely that the L2 writers can be certain about the existence of the geminating vowel /ii/ in this type of adjective as long as they are aware of its word-class. This should be very helpful for L2 writers compared with the situation for another heavy (bimoraic) syllable concerning geminate consonants in a verb. As has been pointed out, knowing the distribution of the presence or absence of a geminate consonant in the gerund-form of a verb depends very much on individual acquisition of grammatical knowledge: i.e. the gerund formation of each verb. Having got these relatively error-free long vowels /aa/, /ii/ and /ee/ out of the way, a close look at more problematic geminating vowels /uu/ and /oo/ at word-final position may be fruitful.

Both writers in the current study seem to have a production problem with /uu/ as they both made a repair for the missing word-final /u/ in /is.sju.u/ (*one week*), and a writer made a repair with omitted /o/ in /to.o.kjo.o/ in the same environment. Two points need to be made. First, a great deal of literature has pointed out that Japanese long vowels tend to shorten in word-final positions. According to Kubozono (2001, p.172), the historically long vowel /oo/ in particular has a tendency toward this position dependent temporal neutralization in bisyllabic S-J (kanji) compounds. Obviously, this temporal neutralization tendency makes L2

---

[7]There is another distinct group of nominal adjective so-called na-type adjectives.

perception of durational contrast difficult. Next, it is also well known in L2 research that English-speaking learners of Japanese have perceptual and production problems in discriminating long vowels from short vowels (Nagai, 1997; Toda, 1998; Hirata, 1999; Oguma, 2001).

While these findings in speech research are very suggestive, we must carefully examine the above cases as written errors. The reason is that deletion errors in my data at the word-final position may be a realization of three potentially different underlying error sources or a mixed influence of various factors. Firstly, it may be that L2 writers had a perceptual discrimination difficulty with regard to the durational contrast, and writers' perceptual domain directly had an impact on orthography. Under this view, written errors are seen as a reflection of incorrect internal representations. Secondly, we cannot deny the possibility that L2 writers somehow managed to hold the correct internal representation despite their perceptual problems; yet some other factors such as imperfect acquisition of orthographic knowledge would have caused the occurrence of these written errors. Thirdly, writers may be able to discriminate the contrast between long and short vowels perceptually, but their lexical encoding is somewhat deficient due to the lack of a comparable lexical contrast in English (L1). This drives us to a question whether deletion errors of a long vowel /oo/ in L2 written production are a result of underlying phonological representation or inadequate orthography acquisition. In order to answer this question, we first need to tap L2 internal representations of Japanese words, especially those that include heavy (bimoraic) syllables /N/, /Q/ or /VV/ where L2 written errors were typically found. After establishing an L2 perceptual baseline in the next chapter on Moraic Awareness, the above question regarding the deletion of long vowel /oo/ in L2 production will be taken up later on in Chapter 6.

However, at this point, one thing should be noted for the development of further experimental materials. In this study, English words were used to induce the sound of target Japanese words by means of meaning. Most of the targets were written as originally intended versions, but there were 3 cases - *gold, house* and *ice*, for which an alternative synonym was recalled. For instance, 黄金 'oogon' was written in place of 金 'kin' and お宅 'otaku' appeared instead of 家 'ie'. In further study, we need to be careful to choose target words in terms of transparency of meaning and availability of synonyms, otherwise we should consider a new means to conduct a similar recall experiment with an increased number of subjects.

4.3 Kanji error monitoring skills

4.3.1 Goal

This section is concerned with kanji production problems encountered by English-speaking learners of Japanese when writing with J-word processors.[8] In the previous sections (4.1 and 4.2), we mainly dealt with hiragana written errors that were made at the initial stage of the writing process in J-word processing. These errors are normally easy to detect, because phonological anomalies of target words can be clearly indicated by hiragana letter(s). Consider if there is a word that corresponds to these hiragana strings. If the illegal hiragana chunks have no corresponding real word, the J-word processor rejects the conversion of the chunks into a kanji, so that the writer either makes repairs and replaces it with the right version or leaves it as an illegal hiragana chunk in the text. In contrast, the writer may end up with having an absolutely irrelevant kanji with perfect shape, if the illegal hiragana chunks correspond to real words (hereafter 'illegal hiragana kanji').

Here, the problem is that current models of J-word processors can offer no meta-information on the checking of wrongly chosen kanji in the final output. The following example demonstrates an L2 writer's problem with regard to this type of kanji conversion error. An illegal hiragana kanji 政界中 'se-i-ka-i-ju-u' (*throughout the political world*) was written for the target word 世界中 'se-ka-i-ju-u' (*all over the world*) just because the writer inserted a vowel /i/ after the first CV sequence /se/. On the surface, this error looks as if it were a homophone kanji conversation error, but actually this illegal hiragana kanji arose from a more complex cause - a hiragana input error involving a vowel insertion was made, then this hiragana chunk was converted into an unintended real kanji, which was followed by a pattern recognition error. It is clear that the less kanji knowledge the learners have, the more problems arise in their visual recognition, leading to more errors as a result of poor self-monitoring. Certainly, this type of kanji error formed a category in my L2 word-processed written error corpus besides familiar homophone kanji conversation errors. These illegal hiragana kanji also create substantial difficulty for readers by forcing them to trace back to the error kanji, which usually indicates a completely out-of-context meaning to the reader and provides them with only partial or no phonological information of the intended word.

---

[8]This section is based on a conference paper by the author (Matsumoto-Sturt, 2002b).

Thus, this study (Experiment 3: Kanji error monitoring skills) aims to observe kanji verification processes among English-speaking users of J-word processors, and to investigate further the meta-cognition problems that lead to L2 learners' error detection failures. Again, samples of concurrent protocol data (Ericsson & Simon, 1979) were recorded from L1 and advanced L2 participants' thoughts while they solved the tasks. Finally, the transcribed protocol data will be analyzed to determine why and when kanji errors occurred.

4.3.2 Method

Participants

L1 participants were both female, and came to Britain as exchange students about seven months prior to the experiment. L2 participants were two undergraduate students, one male and one female, who lived in Japan for about 10 months taking Japanese language courses at host universities. They took part in the experiment for a small payment. The male participant was native speaker of English, and female participant was a German-English bilingual. At the time of the experiment, L2 participants had almost completed their final year courses at the University of Edinburgh: Advanced Japanese oral (discussion and speech presentation), Translation into and from Japanese, and Essay writing in Japanese, which included a word-processed essay (3500 to 5000 Japanese characters). Hence, their authentic language exposure was much longer and wider than that of the novice participants in Experiments 1 and 2.

Materials

60 target kanji or kanji compounds consisting of 2 to 4 morphemes were selected for this task (see Appendices F (material), G (instruction) & H (a sample at actual size). Most of the kanji in this study were taken from my L2 kanji error corpus. Thus, I did not control for the number of morphemes in the target kanji, frequency of each kanji compound, the degree of difficulty, the position of each error in the given target word, nor word-class. 20 items contained no errors and were used as fillers. 20 items were homophone kanji conversation errors, hence containing phonological information of the intended word (henceforth P+). Another 20 items had no cue for phonological information (henceforth P-) with various underlying error sources. For

example, an omitted vowel /o/ resulted in 諸説 /sjo.se.tu/ (*various opinions*) for 小説 /sjo.o.se.tu/ (*novel*); an invented spelling 村料 (there are several underlying sounds of this invented kanji compound) occurred as a suspected copy mistake for 材料 /za.i.rjo.o/ (*ingredients*). These items were further classified according to their visual similarity (V+ or V-) and semantic similarity (M+ or M-). The following table 4.3 shows a typical example of each category:

Table 4.3: Examples of kanji errors and their categories in Experiment 3

Category	Errors	Targets	Error type
P+V+M+	子共 ko-to.mo (child-both)	子供 ko-do.mo (child)	SFM are very close
P+V+M-	弟一 da.i-i.chi (brother-one)	第一 da.i-i.chi (the first)	S & F similarity, but distant M
P+V-M+	中ば na.ka-ba (inside)	半ば na.ka-ba (half)	S and M similarity, but distant F
P+V-M-	寒単 ka.N-ta.N (cold-simple)	簡単 ka.N-ta.N (easy)	S similarity, but distant F and M
P-V+M+	焼やす ya.ki-ya-su (burn)	燃やす mo-ya-su (burn)	F & M similarity, but distant S
P-V+M-	牛後 gyu.u-go (cow-after)	午後 go-go (afternoon)	F similarity, but distant S and M
P-V-M+	寝い ne-i (lie down)	眠い ne.mu-i (sleepy)	S similarity, but distant F and S
P-V-M-	愛増 a.i-zo.o (love-increase)	愛情 a.i-jo.o (love)	SFM are all distant

S=sound, F=form, M=meaning

Next 60 meaningful short sentences or noun phrases were created (P+=20, P-=20, Filler=20) with selected error kanji, and then randomized.

Procedure

All equipment for Experiment 3 is exactly the same as for Experiment 1. Each participant was tested individually in a quiet room. Each sentence was displayed on computer screen one at a time with 18 points size characters (see Appendix H). The participants were asked to click a downward arrow at the lower right of the screen when answered a target, then the next sentence replaces the completed one. The order of experiment is as follows: (1) instructions, (2) practices, and (3) trials. First, the experimenter asked a participant to read instructions, then gave the participant 5 practice trials, including items similar to those in the experiment. These practice trials were also intended to give participants practice for producing desired think-aloud data.

The participants were first asked to make their judgment on whether a given sentence was correct (task 1). Next, they were asked to make a copy of kanji that they thought had been written wrongly, and then write their corrected version of the kanji or kanji compound (task 2). The task 2 was followed immediately after the

task 1 for each target sentence. They were not allowed to use any on-line or off-line aids while they were solving each target, but there was no time limit to complete both tasks.

4.3.3 Results and discussion

Every participant in this study completed the task. However, two instead of four sets of data were analyzed as a consequence of an apparatus problem: Lotus Screen-Cam97 crashed several times during two trials with one L1 and one L2 (German-English bilingual female) participant. The nature of this experiment did not allow them to redo their trials, and the lack of half data (i.e. real-time screen movements) made it impossible to include these sessions for analysis. However, a comparison of L1 and L2 performance was possible with remaining two sets of almost [9] complete data. The following table 4.4 is a summary of their overall performance:

Table 4.4: The number of error detections, correct production of identified errors, and the number of false alarms

Target	Performance	L1	L2
P+	Task 1 (detection)	20 out of 20 (100%)	19 out of 20 (95%)
	Task 2 (production)	18 out of 20 (90%)	15 out of 19 (79%)
	False alarm	0 out of 10 (0%)	2 out of 10 (20%)
P-	Task 1 (detection)	20 out of 20 (100%)	15 out of 20 (75%)
	Task 2 (production)	20 out of 20 (100%)	9 out of 15 (60%)
	False alarm	0 out of 10 (0%)	0 out of 10 (0%)

The L1 participant almost automatically identified the P+ target errors immediately after the targets were displayed. As we can see from the table above, the accuracy of her detection was perfect, though she made two homophone kanji conversion mistakes in her error correction task. We had poor L1 think-aloud protocols from P+ targets, because she made her judgments in a fraction of time without really thinking. However, her task performance average time per item for P- targets, which blocks the phonological root of the target word, was 6 seconds longer than that of P+ items, and produced more L1 protocol data. She commented that difficult items for her judgment were not kanji with a complicated figure, but rather simple kanji that looked somehow wrong to her.

---

[9]As has been mentioned above, a complete set of data such as the raw reaction times of each participant could not obtained, due to trouble with the apparatus. Thus, further chapters will employ a more reliable measurement method in order to avoid the same problem.

Turning now to L2 data. Note that the L2 participant demonstrated a native like error monitoring skill (19 out of 20) for P+ targets, but he failed to reproduce the correct version of identified 4 errors in task 2, which sliced 16% off from his higher detection rate. As indicated by the imbalance of figures in tasks 1 (M=85%) and 2 (M=69.5%) in the L2 data above, the L2 participant in this study had production problems when he could not recall the whole word sound of the target kanji as an input to J-word processor. The data is brought in line with the L1 developmental literature of both oral and written language learning, in which it is claimed that word production on the whole is harder than word recognition (Ehri, 1987, 1992; Henderson, 1982; Treiman, 1993). He also made two false alarms against the correct kanji. Nevertheless the total time that he spent for each session was much longer than those of the L1 participant. For instance, the non-native participant took about 4 times longer (M=68 seconds) than native participant (M=16.2 seconds) for P+ targets. However, more importantly, we sampled plenty of L2 protocol data. The task was to track down the correct kanji from a problematic target, accordingly the L2 participant needed to pay special attention to each item in the given sentence. For this reason, the flow of his thought appeared to become much slower than usual, so the participant managed to keep the instruction - 'voice as you think' fairly well. Let us now look at some L2 protocol samples in detail.

Examples of L2 reading strategies

Example 1 (Prompt decision)

　　　復雑なシステム (fuku̲zatu-na system: 'A complicated system')

1. aaa, it looks ok, system, system.
2. OK, I'm just go'na check *fukuzatu*.
3. [TYPE: fu-ku-za-tu]
4. [the correct version 複雑 appeared on the screen] Ok···aaaa, right.
5. How I checked 復雑, and I've just found out the first character in the original sentence is wrong.
6. So, that's a good job I checked it.
7. The character, which I now write as FUKU of o-o-fu-ku (往復: both ways), and put that in as the incorrect kanji.

This is a particularly straightforward example, which took only 25 seconds. In fact the error was correctly identified as a visually and phonologically similar kanji.

Example 2 (false alarm)

日本の夏は高温多湿だ ('Japanese summer is hot and humid')
nihon no natu wa kooon tasitu da
Japan of summer PP high temperature much humidity COP

1. *nihon no natsu wa* KOO···aaaa
2. OK, aam, I've got a four-character compound.
3. Aaam, trying to think what it is···
4. I don't recognize it.
5. I think ATA in the first two items are not necessary for 'weather'.
6. So, I'm going to change that.   [TYPE: a-ta-ta-ka-i, the converted into 暖かい (warm)]

This is a false alarm that is in fact correct, but he thought it was wrong. From protocol 4, it is clear that he knew neither how to read the compound nor its meaning. Real words are meaningful, hence, easier to hold and operate upon in working memory, but this example shows no sign of being easy to process. In other words, the example indicates the absence of the appropriate lexical entry in the participant's mental lexicon. His strategy here is to divide the unknown kanji compound into 4 separate single kanji and read them in *kun-yomi* (Japanese reading). We notice that he uses his existing knowledge about the different usages between a pair of homophonic kanji 温 and 暖. Although both kanji mean 'warm', their usages are distinctly different. An adjective温かい 'atatakai' (warm) is used for temperature of things, while the other adjective 暖かい 'atatakai' (warm) expresses climate 'warm'. He obviously did not think of another adjective 暑い 'atui' (hot) as a synonym of the half part of this compound 高温 (high temperature), perhaps because it does not share the same sound of a suspicious kanji 温 in *kun*-reading. This example tells us that despite his arriving at this false alarm (a conscious error), he demonstrated his problem-solving skill by creating sub-goals using his existing knowledge, but unfortunately his strategy went wrong when solved his sub-goal.

The L2 participant commented that he managed to comprehend the meaning of a given sentence, because he could often recognize either the first or second constituents of kanji compounds. This implies that he relies on partial recognition of a

given kanji compound not on recognition of kanji as a whole word. That was why it was difficult for him to make judgments about whether each kanji compound in a given sentence corresponded to a real word. It is interesting to compare this remark with Hirose (1992)'s L1 study. In his kanji priming experiments with L1 university participants, Hirose (1992) demonstrated the role of the initial kanji in a given two-kanji compound word in the storage and retrieval of kanji compounds in memory. Hirose claims that the lexicon of kanji compounds was formed according to the meaning of the same first kanji in the L1 mental lexicon. In other words, the meaning of the first kanji in two-kanji compound words constantly plays a role in L1 retrieval of kanji. In the case of L2 kanji reading, if we assume the above L2 comment to represent L2 general reading problems, L2 kanji retrieval processes would be activated in a different way; to be more precise, a utilization of the meaning of the first kanji in two-kanji compound words in L2 retrieval of kanji cannot always be expected. Furthermore, it is reasonable to assume that L2 reading time of single kanji would be different from a two-kanji compound unlike L1 reading, since L2 readers do not always seem to recognize two-kanji compounds as a single meaning unit (word). The following example illustrates another problem of focus of attention, in which the L2 participant finally judged the target error as correct, even though he identified the error in the target kanji compound.

Example 3 (L2 reading strategies)

'There was a <u>ghost story</u> (summit) of Japanese and American leaders in Tokyo'

<div align="center">

東京で日米首脳 怪談 (会談) があった

Tokyo de   niti.bai-shu.noo.-<u>kai</u>.dan   ga atta

Tokyo in Japan.America-leaders-meeting PP COP-past

</div>

1. OK, aaa in Tokyo, American-Japanese leaders at a conference of some kind?'

2. Um, the only kanji that looks out of place to me is the third kanji in compound.

3. So, ke-i⋯.ke-i-da-n. OK. 〔the list shows 軽暖 but not the target 怪談〕

4. Right. Um, so I don't recognize the last kanji in compound that means 'discussion'.

5. It seems to be correct.

6. I'm going to try writing so-o-da-n 〔typing〕 , and see what it gives me OK, aaaa, SOODAN is 'discussion'.

7. This could be the kanji that it 's supposed to be.

8. I'm asking for other versions, now.

9. OK, then. Write in SOODAN and see if it gives me a kanji that is similar to one that is in suggestion.

10. 〔 装弾, 相談 〕 are displayed in the list〕 Aaa⋯.it doesn't look like it.

11. OK, well, it's not coming up under SOO.

12. It's not KEE. I thought what it originally was.

13. So, I can't recognize kanji.

14. Aaa, I try one more thing.

15. I try ka-i. KAIKETU.

16. For some reasons, it should be either KEI or KAI.

17. And see if it gives me a word.

18. 〔 怪談, 会談, 階段, 戒壇 〕 were displayed in the list〕 Right, it does give me 怪談.

19. So, I assume that is correct even though I'm not entirely sure what the meaning is. 怪談 is a word.

20. So, I'll take that.

21. So, 'in Tokyo there was a meeting of Japanese and American leaders'

22. Now, I am just checking 'leaders' 〔typing by saying 'sjo-no' for 'sju-no-o'〕 wondering whether I haven't missed mistakes somewhere else.

23. 〔the desired 首脳 did not appear〕 Um, ok. Aaaa, right. I'm just trying to get the correct reading.

24. Ah, typical problem of long or not vowels.

25. OK, so SHONO. 〔typing by saying 'sjo-no' but the desired 首脳 'sju-no-o' did not appear〕 No!

26. So, I give SHUSHO-san. 〔typing as sju-sjo-o-sa-N〕 Aaaa⋯ No. OK, trying to check NO or NOO is taking time.

27. Right. So, I'm go'na stick with my original. That's correct.

First of all, the participant intuitively identifies the meaning of the correct version of the sentence (the goal at the higher level is correctly reached), but he gets the wrong pronunciation of 'meeting' *kaidan*. Next, his strategy of decomposing the unknown kanji into its semantic and phonological components did not work because the phonological component of the kanji had an irregular reading. At this point, he tried another strategy. He tried to take a known word involving the second kanji in the compound, but this happened to be the wrong word. Then, he tried a third strategy, he seems to realize at this point that the first kanji could have an irregular reading *kai* as well as its regular reading *kei*. But now he loses track of the task, and tries to work out whether the compound is a word or non-word by checking the homophone list given as output by the word processor. Because his attention is now focused on the wrong task, he does not notice the correct version of the compound 会談 in the word processors' homophone list. Again, the analysis of this episode reveals that his step-by-step reading strategies were valid, but because he fixed on a different target, namely checking whether the compound 怪談 is a real word, he ended up with a wrong answer to this particular task.

According to Harada (1999, p.44), we often make errors because of not knowing how to solve the problem; nonetheless we also make errors because we fail to conceive the right move among possible alternatives to solve the problem. Under this view, the participant's difficulty lay in determining what a valid move was, rather than trying to solve the problem itself. I assume that the participant could have identified the right kanji from the list, if his goal was to identify the right kanji to replace the error 怪談 (*ghost story*), rather than to find out whether the kanji compound 怪談 was a real word. The participant clearly said that the meaning of the target kanji was 'discussion' in episode-4. Moreover, the target kanji 会 in 会談 (*talks*) is not a difficult one to recognize for this level of learners of Japanese. Nevertheless, he did not see the correct kanji that was displayed next to the wrong one in the list. In short, his protocol analysis gave us a clear picture of an example of L2 kanji verification process, from which we can see how and why this particular problem-solving behaviour deviates from the ideal conclusion.

4.3.4 Summary of preliminary data

I shall summarize the main points that have been made in this section:

1. The L2 participant faced particular problems at the initial phonological input stage,

2. The L2 participant could not trace back the illegal hiragana kanji (P- targets) more often than homophone kanji errors (P+ targets),

3. The L2 participant showed a tendency to make false-alarms by wrongly classifying a given kanji compound as a non-word,

4. The L2 participant often employed a trial-and-error strategy, whereas L1 participant's judgments were almost automatic,

5. While the L1 participant used both bottom-up and top-down strategies freely in the search for the right kanji, the L2 participant was less able to deal with the task with a holistic approach, and the tasks were locally handled at phonological, semantic or visual level of information of a given kanji.

4.3.5 General discussion and conclusion

This chapter has investigated the L2 writing process of J-word processor through observations and analyses of verbal protocol data in three pilot experiments. Experiment 1 (hiragana copy task) has presented three preliminary findings, which shed light on the L2 specific phenomenon of the appearance of orthographically illegal hiragana strings that are left unconverted into kanji. First of all, L2 novice writers showed visual discrimination problems that were affected by surrounding context (visual recognition problem). Second, L2 writers tended to rewrite more when they wrote a word involving alternative hiragana letters that share the same phonetic value (orthographic problem). Here, the orthographic choice depends on one's conscious use of orthographic conventions. Thus, the writer cannot choose the right one on the basis of phonological knowledge. Third, a word that includes either moraic nasal or geminate consonant tends to trigger a repair because of the need of hitting the same key twice (typing problem). Experiment 1 (hiragana copy task) has shown that orthographically illegal hiragana appear even for the simple copy task where errors were made not only as a result of basic typing mistakes, but also under the influence of more general reading and writing problems. I shall return to this point later.

The production study Experiment 2 (hiragana production task) has presented L2 specific phenomena in three preliminary findings. Firstly, a combination of grammatical formation problems and slips seem to be the main source of errors concerning addition or omission of geminate consonants. But I shall make no further investigation into this; to do so would involve us in a discussion of individual differences in acquisition of grammar in writing production that is of no immediate relevance to the current study. In short, if a further study is conducted to deal with geminate consonants, then it will deal with non-derivational geminate consonants in the noun category. Secondly, L2 writers occasionally decomposed a kanji compound on purpose, and then tried to write individual characters to form what they believed to be their target kanji compound[10]. Thirdly, long vowels /aa/, /ii/ and /ee/ were found to be relatively error-free, while long vowels /uu/ and /oo/ at word-final position may be found to be more problematic for L2 writers. Alongside the findings in previous L1 and L2 research, this part of Experiment 2 has generated a research question about whether deletion errors of the long vowel /oo/ in L2 written production are a result of underlying phonological representation or inadequate orthography acquisition. I shall deal with the question using a set of more controlled experimental materials for a larger number of participants in chapter 6.

One final point is that L2 participants were sometimes unable to remember the whole word sound, though they clearly held the form-meaning connection of a given word. In other words, they can recognize and recall the shape of such kanji, but they cannot write it with J-word processor. Similarly, Experiment 3 (Kanji error monitoring skills) revealed that even the advanced level L2 writer had the same problem. To be more precise, there was an imbalance between in his kanji recognition and production ability, which occurred due to a recall difficulty of the whole word sound of the target kanji. In terms of L2 kanji teaching, findings such as these suggest that the acquisition of the whole word phonology of kanji compound (i.e. words) by means of strong connections of sound-form and sound-meaning is crucial to L2 Japanese electronic literacy. This implies that practice should be focused on vocabulary development rather than rote memorization of each sound and form-meaning relations of individual kanji that are typically employed in the traditional hand-written biased kanji teaching methodology.

Experiment 3 suggests the different efficiency levels of L1 and L2 writers. While both L1 and L2 writers used both bottom-up and top-down strategies, L2 writers

---

[10]Although we cannot completely exclude a possibility of finding a similar L1 writing strategy, we can safely assume that L1 writers seldom write in this way, because this strategy obviously interrupts the flow of overall writing process.

tended to handle tasks locally at phonological, semantic or visual levels of information for any given kanji. This contrasts with the much smoother search kanji search processes of the L1 participants. Further protocol analysis suggests that the L2 participant relied on partial visual recognition of a given kanji compound unlike L1 recognition of kanji as a whole word. This tendency was illustrated by a case of false alarm. Obviously, the participants must have had heavier processing loads than typical word processing activities, because they needed to trace back these heavily distorted L2 errors. Even so, this study presented enough evidence to show the complicated nature of kanji verification processes where more factors other than one's kanji knowledge or lack of attention are likely to play a role. Such factors could be the effects of phonological and visual similarity and contextual relatedness. Certainly, many L1 researchers (Van Orden, 1987; Wydell et al., 1993; Matsunaga, 1995; Sakuma et al., 1998) in the kanji reading literature have paid a great deal of attention to the interactive relation between visual similarity and phonological information in kanji reading, but most of them studied kanji words in isolation, hence excluded information from contextual top-down processing. Here, we should be aware that word-processed written errors almost always appear within a meaningful sentence. Therefore, we ought to consider the contextual effects on the writer's error kanji monitoring process when investigating interactive relation between visual similarity and phonological information in kanji reading. In chapter 7, we shall investigate this issue of the impact of visual similarity and contextual relatedness on the L2 kanji verification process.

CHAPTER 5

MORAIC AWARENESS BY L2 LEARNERS

5.1 Motivation

This chapter is motivated by the findings of the previous chapters. In particular, the Error Analysis in Chapter 3 provided evidence that EJ-specific hiragana errors constituted a major category in our word-processed spelling error corpus. We have good reason to suspect that the structure of the kana syllabary is in some way reflected in L2 word-processed spelling errors in Japanese. Moreover, the hiragana copy task in the preceding chapter has provided some examples leading us to assume that the seemingly easy task of re-typing the given hiragana phrases or sentences required something more than scanning and kana-to-sound unit matching. If our EJ word-processed spelling errors mirror in part the structure of the kana syllabary, this must be because each kana character represents a mora-sized sound unit.[1]

We conducted a preliminary study to examine whether learning kana letters was easy for our absolute beginners, with no previous knowledge of the Japanese writing system. For this, we measured the hiragana achievement of absolute beginners, who are categorized as 'monolingual students', about a month after they attended four 1-hour classes of kana syllabaries teaching. The results showed that all 14 students in the monolingual group achieved A+ (over 90% of 53 full marks; M=51.7857, SD = 1.1387) for the romaji-to-kana mapping section. This result may suggest that the reason for the EJ specific hiragana error category in our L2 error corpus be that the knowledge of letter sounds alone did not equip EJ writers to encode the intended Japanese words perfectly.

---

[1]Some kana letters do not hold this one-to-one correspondence. These are /CjV/ type morae that hold one-to-two (1 mora-2 letters) correspondence: e.g. きゃ ki-ya for /kja/ .

Recent studies in the L1 literacy developmental literature favor this view (see a detailed discussion in 2.3.1). For example, Treiman (1993) provided a study of invented spellings in English, and demonstrated children's reliance on phonology in phonologically accurate misspellings (e.g. *jumpt* for 'jumped'). Her study suggests that the structure of spoken language would provide a key foundation for studies of young L1 learners' literacy development. In short, we focus on one particular aspect of spoken language structure - the mora in Japanese, and we set as the aim of this chapter to investigate the level of perceptual sensitivity to the internal structure of the Japanese syllable among English-speaking (EJ) learners of Japanese at different proficiency levels. Thus, three different proficiency groups of EJ learners will be compared on a 'phonological tapping task', which requires participants to count the number of sound units.

5.2 Background and research hypotheses

In Japanese, the syllable and the mora often overlap, but this is not the case in special syllables. These special syllables are V, CV, or CjV followed by a postvocalic consonant N (nasal coda), Q (the first half of a geminate consonant), or a post-nuclear vowel R (the second half of a long vowel). Here, the awareness of the mora plays a crucial role in speech perception and production.

Several L1 studies (Bertoncini, Floccia, Nazzi, & Mehler, 1995; Inagaki et al., 2000; Mann, 1986b) used Japanese materials to investigate children's sensitivity to syllabic and subsyllabic units (i.e. morae). Moreover, recent psycholinguistic studies on segmentation procedures for speech sounds provided evidence for the role of the mora in Japanese. According to the language-specific hypothesis (see page 32 in Chapter 2), Japanese speakers employ a mora-based listening strategy and segment speech at mora boundaries, as Japanese is widely known as a 'mora-timed' language, while English speakers, who are equipped with a stress-timed procedure, segment continuous speech signals at the onset of strong syllables. However, balanced bilinguals, according to Cutler and Otake (1994), can suppress their native language specific segmentation when the situation demands them to listen to a foreign word, but monolinguals cannot. Under this view, we would expect English monolingual speakers to apply their language specific stress-timed listening procedures to Japanese speech sounds, which do not share with English the property of lexical stress, regardless of their relative applicability. However, note that listeners use many different cues for segmentation. For example, Dupoux et al. (1999)

conducted a cross-linguistic (French and Japanese) perceptual study on Japanese devoicing environments, and they reported that only Japanese listeners predominantly hear illusory epenthetic vowels within consonant clusters even when the vowel [u] was totally removed.

We know from previous research that Japanese listeners segment speech sounds mora by mora, but it does not necessarily follow that Japanese sound units must be segmented only at the level of morae. For example, segmentation at syllable boundaries may also be a viable alternative monitoring strategy for EJ learners, since all syllable boundaries in Japanese are also mora boundaries, though the reverse is not always true (Kubozono, 1989). Moreover, there is evidence that native English speakers prefer to split words before an accented syllable or a stress unit boundary (see Cutler & Young, 1994; Treiman & Zukowski, 1988), and the accent-bearing unit is the syllable in Tokyo Japanese. Syllable-by-syllable segmentation predicts that the syllables are the smallest rhythmic units. Thus the listeners should not further divide syllables into moraic units. In addition, if EJ listeners rely on L1 phonotactics, we expect to observe them to perceive a two-consonant cluster (CCV) in a devoiced CVCV sequence (e.g. 'sukiyaki' as ski-ya-ki), even though such two-consonant clusters are illegal in Japanese.[2] This yields the following research questions:

1. Does the devoicing vowel environment lead to a decrease in the perceived number of rhythmic units? If so, is this decrease consistent amongst the different learning groups?

2. Do English-speaking learners perceive rhythmic units according to a mora-based strategy, or according to a syllable-based strategy? Does this strategy differ according to learning group?

This study provides monolingual baseline data from a group of absolute beginners of Japanese with no previous exposure to Japanese speech sound and the writing system. In addition, we expect to illuminate the way in which more advanced learners of Japanese monitor Japanese words. I predict that English speakers with none or very low exposure to Japanese language employ their L1-language listening strategy

---

[2]The devoicing of the high vowels /i/ and /u/ noticeably occurs between voiceless segments in Tokyo Japanese.

when they are presented with Japanese words, whereas a group of advanced learners of Japanese will have developed near native moraic awareness, and hence are sensitive to the mismatch effect of certain words that include special syllables, and the segmentation patterns will demonstrate their exploitation of a L2 mora-by-mora segmentation strategy. The higher the L2 language competence, the greater will be the learners' level of perceptual sensitivity to the internal structure of the Japanese syllable, resulting in less use of L1 language-specific listening strategies. In short, we seek to determine the extent of the use of L1-specific listening strategies, and in particular, whether EJ learners use the syllable or mora as the basic perceptual unit of segmentation.

5.3   Method

5.3.1   Informants

We tested forty undergraduate students (24 male, 16 female) at the University of Edinburgh. All of the informants in this study were native speakers of English. They were selected to represent largely three groups of English learners of Japanese at different proficiency levels: absolute beginners (i.e. the monolingual group, N=14), literate beginners (i.e. the novice group, N=15), and skilled advanced learners (i.e. the semi-bilingual group, N=11).

A detailed biographical questionnaire (see Appendix I) concerning participants' language background and language experience with Japanese was developed, with two purposes. Firstly, it was aimed to select native speakers of English from the entire class with a size of about forty students in an absolute beginners' class 'Japanese 1' (J1). Secondly, the English monolingual group was separated from other beginners of Japanese who already had a certain degree of Japanese language exposure before entering into higher education. According to their language background, twenty-two English native speakers were identified in J1, and this group was further narrowed down to fourteen as the *monolingual group*, and the remaining eight English-speaking J1 students, who already exposed to Japanese elsewhere, were dropped from the study.

Literate beginners (N=15) in their second year formed the *novice group*, with approximately 150 hours of classroom exposure at the time of testing. Most members of the novice group had never lived in Japan. They could write and recognize

about 180 kanji in addition to their kana literacy skills (i.e. ability to read and write hiragana and katakana syllabaries). At the commencement of the 2nd year course, I normally rate second-year students' spoken ability in a novice proficiency range when their oral skills are assessed by the standardized oral proficiency testing method of ACTFL OPI.[3] This means that they can only convey minimal meaning using isolated words and rote phrases, and their pronunciation may be strongly influenced by their mother tongue; thus, they often end up with misunderstandings. I chose 4 students at random from this particular group and conducted OPI individually, all of which were rated in the novice range.

Eleven fourth-year students, all of whom had experienced a period of 10 months study program in Japan, formed the third group, the *semi-bilingual* group. Their literacy level was at an advanced level. For example, one of their course assessments is to write an academic essay in Japanese on a given topic that covers political, environmental, historical or cultural issues. Again I randomly chose 4 students from this particular group and tested their oral proficiency level, all of which were rated in an advanced range. According to the ACTFL guidelines (ACTFL, 1999, pp.83-85), speakers at the advanced level can handle a great deal of communicative activities relating to work, school, home and leisure situations with linguistic ease and confidence. From the viewpoint of native listeners, advanced speakers speech was understood by a native speaker without any difficulty. Importantly, the criterion of 'native speaker' excludes experienced foreign language teachers, thus, the native listeners here means ordinary speakers of a given language who are not used to dealing with a non-native speaker's spoken language. Although they are fluent communicators, this level of speakers tends to show linguistic weakness when a more complex situation demands them to discuss abstract topics that require special fields of expertise and the use of precise vocabulary. This means that they are comfortably accurate and fluent as long as they are discussing domain-specific concrete topics. For this reason, I do not regard them as perfect bilinguals, but well-developed semi-bilingual speakers of Japanese.

5.3.2 Materials

To test EJ learners' sensitivity towards either the syllable or the mora in each word, we adapted the much-used 'phonological tapping task' (see Inagaki et al. (2000),

---

[3]ACTFL (American Council On The Teaching Of Foreign Language) OPI (Oral Proficiency Interview).

Mann (1986a)) or 'counting test' (test materials, see Appendix J). This task has been used to test phonological awareness in L1 studies, where young children are typically asked to tap out the number of phonemes or syllables in each word by jumping a doll or tapping the number with a hammer, while they were segmenting the given words. However, we developed a new method to replace the 'jumping a doll' segmentation task, because we are dealing with adult participants rather than children. Informants were asked to make an intuitive judgment about how many sound units or rhythmic beats they perceived every time they heard a Japanese word, and the participants were required to circle the number that they perceive to correspond to the number of rhythmic units (see Appendix K). Altogether 80 Japanese words were selected for the test, half of which were 3-syllable 4-mora words and the other half materials consisted of twenty ordinary CVCVCV words (3-syllable, 3-mora) and twenty CVCVCVCV words (4-syllable, 4-mora). Each selected word was strictly controlled so as the following criteria were met:

1. No more than one special (2-mora) syllable per word
2. Except for the 10 items with a diphthongal vowel[4] other items did not contain any diphthongs
3. Except for the 20 items with a devoiced vowel, other items only contained voiced vowels.

(1) included words with the following structures: CVN (e.g. buNgaku *literature*, N=10), CVQ (e.g. saQporo *Sapporo*, N=10), and CVV (long vowels: (R), siRtake *shiitake mushroom*, N=10; and diphthongs, e.g. higaeri *day trip*, N=10). In the CVN and CVR words, the location of the target mora was balanced equally between word medial (e.g. seNtaku *laundry*) and word final positions (e.g. kokumiN *nation*). Similarly, half of the dipthongs in (2) straddled a morpheme boundary (e.g. shima.uma *zebra*) while the other half did not (e.g. uo.nome *corn*). For (3), the devoiced and non-devoiced materials were evenly distributed between the 3-syllable, 3-mora materials and the 4-syllable, 4-mora materials. I included devoiced vowels because the devoicing of the high vowels /i/ and /u/ noticeably occurs between voiceless segments in Tokyo Japanese (Tsuchida, 2001), and we expect to observe EJ participants' perception of a phonologically illegal two-consonant cluster (CCV) in a devoiced CVCV sequence. For example, the CVCVCVCV word 'sukiyaki' might be perceived as 3 syllables CCV-CV-CV ski-ya-ki, because its first

---

[4]In Japanese, it is generally assumed that VV sequences (e.g., /ai/, /oi/) with falling sonority (i.e. falling diphthongs) constitute diphthongs. However, by 'diphthong' we refer more generally to "a vowel where there is a single (perceptual) noticeable change in quality during a syllable" (see Crystal (1991, p.105)).

vowel is typically devoiced, but such a consonant cluster cannot constitute a mora in Japanese.

As far as possible, the study used non-inflected words (N=79), to avoid confounds with grammatical factors. Moreover, familiar words and unfamiliar words were balanced, with 41 items selected from the course textbook, and 39 chosen through exhaustive dictionary search. Finally, the pronunciation and the pitch pattern of all selected items were checked against *Shinmeikai Nihongo Accent Dictionary* before the recording took place. A female native speaker with a Tokyo accent read the randomized list of stimuli in isolation three times, at 10 second intervals, for each word in a soundproof booth. These stimuli were recorded onto conventional audiocassette tape for the classroom use.

5.3.3 Procedure

The tape-recorded test items were played back to the participants. They were instructed to make an intuitive judgment upon hearing them as to how many sound units they perceived in a given word. Participants were provided with an answer sheet and asked to circle one and only one answer (i.e. choosing one number from 1 to 5) per word. I emphasized that the test was not aimed to judge their linguistic knowledge, so that there were no right or wrong answers but the important thing was to trust their own judgment. A practice session was conducted which was followed by the test where the whole class heard each target three times. A short break was inserted after each twenty words were judged in order to prevent them from tiredness or boredom.

5.4   Results

5.4.1 Baseline data

Analyses were carried out to address two separate research questions. The first question was whether or not the devoicing vowel environment led to a decrease in the perceived number of syllables, and whether or not this decrease was consistent amongst the different learning groups. The second question was whether participants perceived rhythmic units according to a mora-based strategy, or according to a syllable-based strategy, and whether this strategy differed according to learning

group. Each of these questions will be answered by performing analyses on a subset of the experimental conditions.

The first research question required looking at devoicing vowel environments. There were 40 items in which the number of morae and the number of syllables were the same. Of these, 20 had 3 syllables and 3 morae (3s-3m condition), and 20 had 4 syllables and 4 morae (4s-4m condition). Half of each of these groups included a devoiced vowel (v-) and half did not (v+). There should be very little error in counting judgments for the v+ experimental items, since both a mora-based counting strategy and a syllable-based strategy will yield the correct result (i.e. a count of 3 or 4 respectively for the 3-3 and 4-4 conditions). However, if devoicing vowel environments cause a decrease in perceived numbers of rhythmic units, these conditions should have lower mean count values than the v+ conditions. Moreover, if this difference between v+ and v- differs according to learning group, we will expect a statistical interaction between voicing status and learner group.

The mean perceived rhythmic units for the 3-3 and 4-4 conditions were submitted to 3-way analyses of variance.[5] These analyses were run both on the data for subjects collapsed over items (F1) and on the data for items collapsed over subjects (F2). Learning group (monolingual vs. novice vs. semi-bilingual)[6] was treated as a between-subjects but within-items factor. Mora count (3 vs. 4) was treated as a within- subjects but between-items factor. Voicing status was treated as a within-subjects, but between-items factor. Thus the designs were mixed 3x2x2 ANOVAs.

The left panel of Diagram 5.1 shows the results for the v+ items for the three learning groups and two mora lengths, and the right panel shows the results for the v- items. The pattern of means suggests that there were very few errors in the v+ conditions for any of the learning groups, but that the perceived counts for the v- items were lower for the less experienced learners. In general, the less proficient group perceived fewer rhythmic units than more experienced group due to the difference in v- (devoiced) counting.

The analyses of variance confirmed that this pattern was reliable, with a significant interaction between subject group and vowel type ($F1(2, 37) = 23.262$, MSe = 0.0146, P <.05; $F2(2,72)=25.779$, MSe = 0.0101; P <.05). There were also main effects of subject group, voicing status and mora count (all p's <.05), and, in the

---

[5]Several non-parametric tests (Kruskal-Wallis and Friedman Chi-Square test) were also run, which showed similar results to the parametric tests. However, because of the possibility of analyzing multi-factor designs, I report ANOVAs.

[6]These categories of learners are defined in the *Informants* section; see pages 90-91.

Figure 5.1: Perceived rhythmic units   Figure 5.2: Perceived rhythmic units
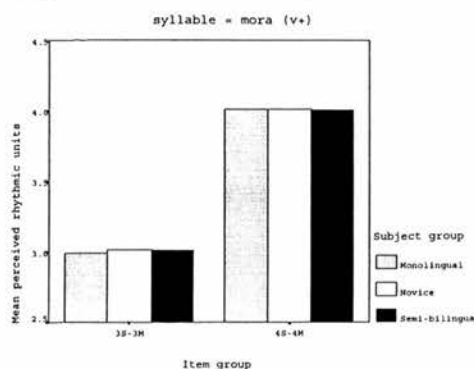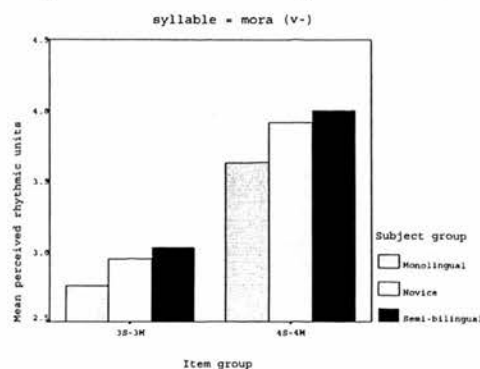
subjects analysis, there was a two-way interaction between mora count and voicing status ($F1(2,37)$ = 8.15, Mse =0.006, P <.05; $F2(1,36)$ = 1.57, Mse = 0.0219, P >.2). The trend was for the difference between v+ and v- to be larger for the 4-4 conditions than the equivalent difference for the 3-3 conditions. We will discuss this finding below. There were no other significant effects in the overall analysis.

Analyzing the data for each of the learner groups separately revealed that there was a main effect of voicing status for the monolingual group ($F1(1,13)$ = 38.65, Mse = 0.005, P <.05; $F2(1,36)$ = 28.97, Mse = 0.033, P <.05) and for the novice group ($F1(1,13)$ = 19.19, Mse = 0.034, P <.05; $F2(1,36)$ = 10.65, Mse = 0.06). Both of these main effects resulted from fewer perceived rhythmic units for the v- items than the v+ items, suggesting that morae with voiceless vowels tended to be perceived as syllables less often than morae without voiceless vowels. However, there was no sign at all of an effect of mora count for the semi-bilingual group (both F's <1), suggesting that morae with voiced and voiceless vowels were perceived identically for this group. For all groups there was a main effect of mora length (all p's <.05), confirming that 3-syllable, 3-mora words were perceived as having fewer syllables in general than 4-syllable 4-mora words. However there were no interactions between mora length and voicing status, except in the subject analysis of the monolingual group ($F1(1,13)$ = 6.95, Mse = 0.010, P <.05; $F2(1,36)$ = 1.57, Mse = 0.033, P >.05). This interaction mirrors the two-way interaction in the overall analysis; the difference between the v+ and v- conditions for the 4-syllable 4-mora items (4.01 vs. 3.63) was greater than the corresponding difference for the 3-syllable 3-mora items (2.99 vs. 2.76). This could suggest that the four-mora items were particularly difficult to perceive for the monolingual group when a voiceless vowel was involved, perhaps because of perceptual overload from the increased number of

syllables coupled with the devoiced vowel environment.

5.4.2 The perceptual units of three subject groups

The data for the second question (i.e. mora-based or syllable-based strategy) were submitted to a two-way analysis of variance (ANOVA) with the between-subjects but within-items variable of learning groups (monolingual vs. novice vs. semi-bilingual) and the within-subjects but between items factor of item group, with four levels: N (nasal coda), Q (the first half of a geminate consonant), R (the second half of a long vowel), and dip (diphthong). Each item-group had 10 items of 3-syllable 4-mora words. Diagram 5.3 below shows the results for the three learning groups and four conditions. The pattern of means suggests that perceived rhythmic units are generally fewer on all items for the less experienced learners than more experienced groups, though the difference between monolingual and novice groups are not as big as the difference with semi-bilingual group.

Figure 5.3: Perceived rhythmic units
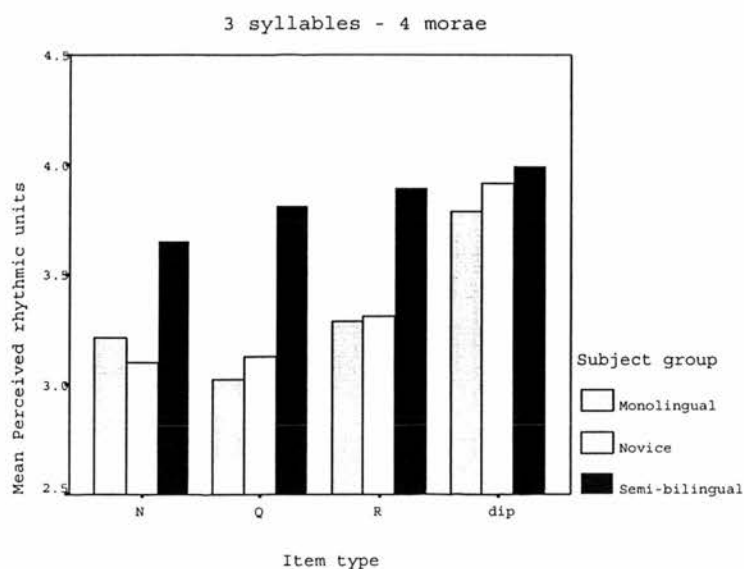
To confirm these findings, analyses of variance (2 factors mixed design) demonstrated a significant interaction between subject group and item type ($F1(6, 111)$ = 7.203, Mse = 0.0415, P <.05; $F2(6,72)$ = 16.00, Mse = 0.0151; P <.05). There were also main effects of subject group and item type (all p's <.05). The pattern of means suggests that the interaction occurred because the semi-bilingual group used

a mora-based strategy (resulting in 4 perceived rhythmic units) relatively often for all item groups, but the other two learner groups differed more in their strategies according to item group. In general, the diphthong materials were perceived to have 4 rhythmic units relatively often for all groups, while there was more of a difference between learner groups for the Q materials and R materials. The differences between item groups were relatively small between the monolingual and novice groups, while much larger differences were found between these two groups and the semi-bilinguals.

We followed up this overall analysis with separate one-way ANOVAs for each learner group, with post-hoc tests to allow pairwise comparisons between item groups for each of the subject groups (Tukey's Honestly Significant Different test (HSD) test). The one-way ANOVAs and pairwise comparisons included the results not only for the R, Q, Dip and N conditions, but also for the v+ versions of the 3-syllable 3-mora items, as well as the v+ 4-syllable 4-mora items. These two latter conditions served as baselines against which the other conditions could be compared. There was a main effect of item type for all groups: monolingual group ($F1(5,65) = 110.077$, MSe $=0.0221$, P $<.05$; $F2(5,54) = 75.579$, MSe $=0.0229$, P $<.05$), the novice group ($F1(5,70) = 78.428$, MSe $= 0.0365$, P $<.05$; $F2(5,54) = 152.074$, MSe $= 0.0125$, P $<.05$), and the semi-bilingual group ($F1(5,50) = 36.953$, MSe $= 0.042$, P $<.05$; $F2(5,54) =241.762$, MSe $= 0.0058$, P $<.05$). Three tables below summarize the results of three multiple-range tests:

Table 5.1: Monolingual group

	3-3 v+	R	Q	Dip	N	4-4- v+
3-3 v+	-	-	-	-	-	-
R	**	-	-	-	-	-
Q	n.s/n.s	**	-	-	-	-
Dip	**	**	**	-	-	-
N	**	n.s/n.s	n.s./ *	**	-	-
4-4v+	**	**	**	**	**	-

'*' denotes significant result; 'n.s' means not significant. The left shows item analysis; the right displays subject analysis

Unexpectedly, the monolingual group (M=3.2143, SD=0.1703) significantly outperformed novice group (M=3.10, SD=0.1773) on N (CVN) condition; the monolingual group perceived the N items to have significantly more rhythmic units than the 3-3 v+ baseline, while the novice group did not. The result of Q vs. 3-3 v+

Table 5.2: Novice group

	3-3 v+	R	Q	Dip	N	4-4- v+
3-3 v+	-	-	-	-	-	-
R	**	-	-	-	-	-
Q	n.s/n.s	*/n.s	-	-	-	-
Dip	**	**	**	-	-	-
N	n.s/n.s	**	n.s./n.s.	**	-	-
4-4v+	**	**	**	n.s./n.s.	**	-

'*' denotes significant result; 'n.s' means not significant. The left shows item analysis; the right displays subject analysis

Table 5.3: Semi-bilingual group

	3-3 v+	R	Q	Dip	N	4-4- v+
3-3 v+	-	-	-	-	-	-
R	**	-	-	-	-	-
Q	**	n.s/n.s	-	-	-	-
Dip	**	n.s/n.s	*/n.s	-	-	-
N	**	*/n.s	*/n.s	**	-	-
4-4v+	**	*/n.s	*/n.s	n.s/n.s	**	-

'*' denotes significant result; 'n.s' means not significant. The left shows item analysis; the right displays subject analysis

pairwise comparison demonstrates that the semi-bilingual group perceived the Q-items to have significantly more rhythmic units than the 3-3 baseline controls, but no clear difference was seen in monolingual group and novice group, presumably indicating a purely syllable-based counting strategy for geminates.

However, in the CVV (diphthong) condition, less experienced learners did show good performance compared to the other three conditions. This means that they often successfully perceived two distinctive morae (i.e. vowels) rather than a sequence of two consecutive vowels within the same syllable. The reason may be that unlike English diphthongs, the first vowel and the second vowel in a Japanese diphthongal VV sequence only change in quality, while length and loudness remain unchanged, so that EJ learners were aware that the second part of a vowel sequence had a stronger intensity than their expectation for English counterparts, in which the second vowel normally has rather shorter and quieter characteristics than the first vowel (Roach, 1983). Finally, the semi-bilingual group showed significantly more perceived rhythmic units for all 3-syllable 4-mora items compared with the 3-syllable 3-mora baseline. Moreover, for this group, the perception of diphthong items was statistically indistinguishable from 4-syllable items. These results demonstrate that semi-bilingual group often used a mora-based strategy for CVN, CVQ,

and CVV items, whereas both monolingual group and novice group did not clearly show great sensitivity to morae, thus largely confirming their use of a syllable-based strategy.

5.5 Discussion

We set out to answer two questions about EJ learners' perception of Japanese words, and the following discussion provides answers to these questions. Our first question concerned the devoicing effect on the perception of Japanese words amongst the different learning groups. The result of the CV baseline analysis revealed that the perceived counts for the devoicing vowel (v-) items were consistently lower for the less experienced learners, while there was no difference for semi-bilingual group. As expected, the monolingual group showed a decrease in the perceived number of rhythmic units for the perceived two-consonant cluster (CCV) in a devoiced CVCV sequence. Surprisingly, even the kana-literate novice group showed a similar tendency, though to a lesser degree. This result demonstrates that the phonotactics of English had an influence on the less proficient learners' perception of Japanese sounds. The novice group still showed evidence for a monolingual listening strategy even after they had completely acquired the CV-based kana syllabaries. Thus the result indicates that the linguistic knowledge of grapheme-to-mora matching was not enough to facilitate moraic awareness for the novice group, perhaps because their linguistic experience was only limited to classroom activities, and they certainly needed more exposure to Japanese sound and print for their sensitivity to mora to kick in.

However, the semi-bilingual group performed significantly different from the less proficient groups by exploiting a typical L1 mora-by-mora segmentation strategy, resulting in their consistently counting both v+ and v- items equally. It is interesting to compare our results with Dupoux et al. (1999)'s study, in which Japanese listeners constantly hear illusory epenthetic vowels within consonant clusters even when the vowel was completely removed. They claim (1999: p.11) that this epenthesis effect is robust because it was still significantly present even when the Japanese participants were experimentally aided to discriminate an epenthesis contrast (e.g. ebuzo-ebzo), and the participants' proficiency in English helped very little to change this pattern. Nonetheless, the semi-bilingual group in the present study gave an extra count to v- items, indicating that they managed to break up consonant clusters at the perceptual level as if they had Japanese ears. Here, we might suspect

that the semi-bilingual group somewhat relied on their orthographic knowledge of CV-based kana syllabaries. According to Kubozono (1995, p.151), this could be understood in terms of their orthographic knowledge (in this case, CV-based kana syllabaries) being 'integrated into the phonological competence' of semi-bilingual group so that they routinely exploited a mora-by-mora segmentation strategy under the indirect influence of the writing system. There is much justice in this view as the phonological awareness literature generally considers the level of learners' literacy as the key factor to determine the performance of the higher-level phonological awareness tasks such as phoneme counting tasks. Thus, we may say that a major factor responsible for the semi-bilingual participants' performance in the mora counting task would be the high level of proficiency in reading and writing in kana orthography. In any case, it is clear that the semi-bilingual group showed sensitivity to Japanese, rather than English, phonotactics at the perceptual level, while the monolingual and novice groups employed an L1-language strategy when they were presented with devoicing Japanese items. This finding provides new evidence to supplement the previous Japanese pedagogy literature, which largely lacks kana acquisition research.

We attempted to answer the question of whether EJ learners perceived rhythmic units according to a mora-based strategy, or according to a syllable-based strategy, and whether this strategy differed at different proficiency levels of Japanese. The results for the special syllable items confirmed that learners relied to a large extent on a syllable-based strategy, but that the extent of this differed according to learner group. For the novice learners, whose data appear in Table 5.2, perception of 4-mora 3-syllable special syllables did not differ from 3-mora CV (v+) baseline controls, except for the CVR items (diphthongs and long vowels). Firstly, although the novice group's mean number of perceived CVQ items was numerically greater than the monolingual group, ANOVA results demonstrated that neither group treated CVQ items differently from the baseline CV (v+). Secondly, both monolingual and novice groups treated CVR items significantly differently from the baseline CV(v+), but their perceived rhythmic units were much fewer (monolingual: M = 3.2857, SD = 0.2770; novice: M = 3.3067, SD = 0.3595) than the semi-bilingual group (M=3.8909, SD=0.0495). Again, this trend was seen in the results of CVN items. In this category, the monolingual group performed reliably better than novice group relative to the baseline, though the actual numerical difference was very small (monolingual: M=3.21, SD = 0.17; novice: M = 3.10, SD = 0.18). These results indicate that novice group's sensitivity to mora is not much different from these of monolingual group, and in some cases perhaps even worse. It is reasonable to assume

that the syllable is the smallest rhythmic units for both of these groups, which is why they are very insensitive to the presence or absence of special syllables (i.e. CVN, CVQ or CVR) in words. Thus, the present results provide further evidence that special syllables, which create syllable-mora mismatch effect, constitute a major challenge for learners of Japanese.

In summary, the different patterns of participants' perceived units confirmed our hypotheses that L1-specific listening effect is strong among English speakers with none or very low exposure to Japanese language, while more experienced fluent learners' moraic awareness is found to be near native level. These semi-bilingual learners were able not only to inhibit or 'switch-off' their L1-specific listening strategy, but also to 'switched on' an L2 mora-based procedure when they heard Japanese sounds. In addition, this study has produced interesting results that suggest a need of further study. For example, we have seen in our written production studies that novice writers as well as semi-bilingual writers had writing problems with the long vowels, but the present study demonstrated that all three groups significantly differentiated CVR items from relevant CV+ control items. One explanation may be that their perception and production abilities are rather asymmetrical. Indeed, some studies, which deal with the relationship between speech production and perception, suggest that the relation between L2 learners' production and perception abilities may not be as symmetrical as we would probably expect (Neufeld, 1980, 1988; Stager & Werker, 1997). In addition, it is also plausible that their difficulty lies at different levels of phonological units, or at other linguistic levels (i.e. orthography). If so, it is interesting to examine the issue of the relationship between perception and production, especially with the same learners' speech and hiragana written production, treating the present finding as a perceptual baseline.

CHAPTER 6

READING AND SPELLING OF JAPANESE WORDS

6.1 Motivation and Goal

This chapter will explore the general question of whether an L2 writer's linguistic problems are reflected in their written products, and we will be especially concerned with the nature of the relationship between phonological and orthographical representations. The tasks in 6.6.1 and 6.6.3 are intended to investigate this question, from which we can expect to assess the transitional aspect of interlanguage (IL) lexical representations, such as patterns of regularity and acquisition order. The spelling task also aims to address the issue of L1 transfer, which has been one of the important areas in SLA research. We are especially concerned with the effect of L1 script (romaji) on L2 writing by English-speaking learners of Japanese (see more detail in 6.3). The naming task (6.6.2) seeks to determine the exact nature of the quality of lexical representation of a given word.

Previous chapters have so far confirmed that some EJ written errors strongly indicate a cross-linguistic problem in perceiving and producing the durational contrast of Japanese words. Much of the literature on English-speaking children's L1 reading acquisition claims that children's knowledge of the spoken language plays an important role in their acquisition of printed language (Gough, Ehri, & Treiman, 1992; Templeton & Bear, 1992). With regard to the issue of the relationship between spoken and written forms of Japanese words, Chapter 5 investigated EJ learners' sensitivity towards the mora as a perceptual unit, in listening to Japanese words. Chapter 5 also provided further evidence that special syllables create a major challenge for novice learners of Japanese. Similarly, we have seen in our written production studies in sections 4.2 and 4.3 that novice writers as well as semi-bilingual

writers had writing problems with words containing the long consonants (geminated) and long vowels. What we do not know at this stage is how phonological and orthographical representations are related in the L2 mental lexicon. A problem here is that all written errors in Chapter 3 were taken from free writing, and hence from observational, uncontrolled data. Thus, the following sections will systematically investigate the perception and production of selected Japanese words.

In addition, the present chapter investigates an issue of spelling (hiragana) ambiguity in Japanese, since some EJ spelling errors point to an influence of Japanese orthography that is based on confusion between two-to-one hiragana-sound correspondences. In the L1 reading and spelling study of Goswami and Bryant (1990, pp.56-57), it was found that exception words (e.g. 'said') and ambiguous words (e.g. 'beef') are hard to spell for children. Their finding was that children made more spelling mistakes when they tried to write ambiguous words than they did with predictable regular words (e.g.'dish') though there was no difference between these words in their reading performance. For example, *beef* is ambiguous with three spelling options: 'ee', 'ie', and 'ea' for the long 'e' sound. If this is also the case for L2 spelling in Japanese, the spelling of words that involve two-to-one hiragana-sound correspondences such as the long vowel /o:/ and the CV sequences of /zu/ and /du/, which share the same phonetic value but are written with two different hiragana letters, are likely to cause problems for EJ learners. Notably, most of the content words are normally written in kanji rather than in hiragana, so that the spelling options for the long vowel /o:/ will not be a nuisance as long as the writer chooses to write the kanji version of these words by hand. In other words, this particular orthographic ambiguity is hiragana specific, thus the writer's orthographic underspecification is masked if such words are written in kanji. This may be the reason that the spelling (hiragana) ambiguity of some words has not received much attention from teachers and researchers in Japanese pedagogy up until now. Nevertheless, we have seen that mastering accurate hiragana orthography is a key to success in J-word processing, so that the issue is absolutely relevant and important to the current study.

One final point should be made about a methodological issue. In order to test accuracy and speed of word pronunciation, we need to consider the frequency of the words, as this is a fundamental factor in word recognition (see more detailed account in section 2.3.3). For example, high frequency words are pronounced faster and more accurately than low frequency words. This created a methodological problem for the present study. To the best of my knowledge, there is no previous

literature in L2 acquisition of Japanese that deals with frequency issues of kanji compound words, so that I needed to consider what kind of database was suitable for the purpose of the present study. One possibility was to consult an available L1 database, which is based on word frequency in ninety Japanese magazines (Kokuritu Kokugo Kenkyuusho 1997). However, it is difficult to be convinced that the L1 data is relevant to our literate beginners, whose language exposure is limited to classroom learning (i.e. tutored acquisition). I assumed that their lexicon is organized differently, as their input is not naturalistic, as most of the input comes from course textbooks and related classroom activities.[1] Accordingly, I decided to build my own frequency database (see 6.4.2). Thus, this chapter employs an original methodology to investigate L2 lexical reading and spelling of Japanese words, from which I expect to find answers to my research questions, as well as providing a new insight into L2 reading and writing research.

6.2 Background

6.2.1 Syllabeme languages

The previous chapter 5 confirmed that the syllable is largely the smallest rhythmic unit for monolingual and novice learners, though their listening strategies differed according to items. On the other hand, the semi-bilingual learners used a mora-based strategy relatively often for all items. It is interesting to see if this listening strategy has any impact on their acquisition of Japanese words. It is possible that the lower proficient writers less clearly distinguish bimoraic syllables from CV-based monomoraic syllables. We can actually observe this phenomenon in so-called 'syllabeme' dialects: i.e. a syllable-based dialect that is used in some parts of northern and southern Japan (Kubozono, 1999; Shibatani, 1990). According to Kubozono (1999), diphthongs are often monophthongized with compensatory vowel lengthening in the mora-based Tokyo dialect, while monophthongization of diphthong does not go along with lengthening of vowel in the syllable-based Kagoshima dialect. For example, 'daikon' *radish* could be pronounced as 'deekon' in casual speech of Tokyo dialect, whereas it is pronounced as 'dekon' in Kagoshima dialect. The reason for

---

[1]See recent relevant articles including N.Ellis (2002), Gass and Mackey (2002), and Tarone (2002) in *Studies in Second Language Acquisition* (2002, 24) on frequency effects in L2 language use.

this difference is that the mora-based Tokyo dialect is sensitive towards the 'preservation of mora', while syllable-based Kagoshima dialect is not concerned with this segmental loss (Hayes, 1989; Kubozono, 1999).

Under this view, it could be possible to categorize the state of the novice writers' interlanguage (IL) as the one that is like a syllabeme dialect, whilst the interlanguage of the semi-bilingual writers could be very close to mora-based Tokyo dialect. For instance, a sample such as the vowel shortening error /sjosetu/ for /sjoosetu/ 'novel' taken from our error corpus, may be a reflection of the writer's syllable-based IL lexicon, where a loss of a mora never affects the syllable structure of this word: (i.e. 'sho.se.tu' or 'shoo.se.tu'). Alternatively, we could interpret the same vowel shortening error as indicating either the incomplete acquisition of phonemic contrast, insufficient knowledge of the orthographic representation of a given word, or a combination of both. In any case, the quality of EJ learners' lexical representations in relation to spoken and written production accuracy would be the key to examine the lengthening or the shortening phenomenon of the long vowel /o:/.

6.2.2 Interlanguage lexical representations

Perfetti (1992, p.146) made it clear that 'reading' entails "the form of knowledge that allows recognition", and he believes that this knowledge or representation is utilized for both spelling and reading throughout the whole course of development. While the necessity of knowledge of a word's orthography in reading and spelling is widely accepted, opinions and evidence remain inconclusive about whether reading and spelling share a single lexicon (e.g. (Ehri, 1987; Perfetti, 1992)) or two separate lexicons (e.g., (Weekes & Coltheart, 1996)) of reading (i.e.perception) and orthography (i.e.production). Regardless of whether the representation of lexical knowledge is shared or separate, there is a fairly general agreement that spelling or reading problems reflect deficits in lexical representations (Ehri, 1980; Henderson, 1980; Kondo, 1998; Stanovich, 1988; Perfetti, 1992; Wade-Woolley & Siegel, 1997).

The idea put forward by L1 developmental models of lexical representations (e.g. (Ehri, 1980; Henderson, 1980; Perfetti, 1992) can be applied to study of L2 acquisition of Japanese lexical representation. The main reason is that this model could handle the learner's ever-changing transitional system (i.e. Interlanguage) containing less specified or gray acquisition areas of even readers of high ability. Perfetti's

Restricted-Interactive Model (Perfetti, 1992, p.163) assumes that the reading lexicon contains two sublexicons: (1) "a developing functional lexicon with representations underspecified", and (2) "an autonomous lexicon with representations fully specified and redundant". A given word increases its precision in the developing functional lexicon until it reaches the autonomous lexicon with fully specified representation, and this process essentially takes place word-by-word. Although Perfetti accepts that measuring an individual reader's representation quality is not so simple, he suggests that we could use spelling as an index. In this model, the acquisition of a functional representation system involves increases in the number of orthographically addressable lexical entries (i.e. vocabulary growth) and increases in the 'quality' of the lexical representations. According to Perfetti (1992, p.157), gaining 'precision' and 'redundancy' means the increase in quality, and partially specified 'variable' representations are replaced with the fully specified superior one. The variable representations mean a skeletal form of a given word where the representation is unstable and changeable. Note that although the variable representations are incomplete representations, it does not necessarily entail inclusion of incorrect letters. For example, there is a hole (i.e. a missing letter) in an underspecified representation of the word 'ir_n' (iron). Perfetti calls the missing vowel a "free variable". In a longer word such as 'tongue', a very early acquisition stage of an unstable representation of this word may have more than one hole. For example, both the vowel letter and the sequence of letters following the n are unidentified in the representation of 't*g*'. Perfetti also illustrates hypothetical 'reversal' (1992: p.158) representations of 'ukulele' as 'uk*l*l*' and 'ukil*': the former is a more specified representation, since the latter represents a "precise but incorrect letter" in the vowel position.

Perfetti (1992, p.157) points out that we can predict certain 'nonarbitrary choices' in the variable representation. For example, in the case of English children, according to Perfetti, it is well known that novice writers include initial letters, thus word medial and final positions are more likely to have free variables in their lexical representations. In addition, English vowels are more likely to be variables than are consonants, because vowels in English have much phonemic variability. By comparison, Japanese vowels have less complicated sound-letter mappings and there is not much literature on L1 hiragana errors. However, there is an example in Kubozono (1998, p.138) in which a beginning hiragana writer at a lower-grade primary school wrote a sentence おとさん o-to-sa-n (*Dad*) がんばて ga-n-ba-te (*do your best*) ね ne (*OK?*), in which a vowel letter for the second half of the long vowel in father (おとうさん o-to-R-sa-N) and a small hiragana っ for the first half

of the geminate consonant in hold on (がんばって ga-N-ba-Q-te) are missing. The answer to the question of whether this writer's phonological representations are fully specified or variable is beyond the current study, but it is clear that this L1 writer's orthographic representations include free variables that are two instances of bimoraic syllables CVR and CVQ, but both word final and the word medial CVN were present.

In addition to this, our production study in Chapter 4 revealed that bimoraic syllables generated similar writing problems among EJ writers for long vowels and geminate consonants. On the contrary, moraic nasal /N/ errors were found to be context sensitive, where nonarbitrary choice for /N/ is predictable from the different phonotactic behaviour between L1 (English) and L2 (Japanese). There is no geminate nasal in English, while Japanese contrast /N/ with the non-moraic nasal consonant, so that if the following segment is a non-moraic nasal (e.g. /N-nV/ sequence) two segments are often reduced into one segment in L2 perception and production. For example, 'ho-N-ne' (*true intention*) is likely to become 'ho-ne' (*bone*). Similarly, a /(C)V-N-V/ sequence such as 'u-N-e-e' (*management*) is a difficult word for EJ learners because of the appearance of the nasalized vowel that is articulately influenced by the adjacent vowel. Perceptually, 4-mora word 'u-N-e-e' (*management*) could be reduced into /Ne:/ for EJ learners' ears. However, our perception study revealed that the English monolingual group differentiated CVN items from the baseline items beyond chance. A reason may be the distributional fact that an open-syllable language like Japanese has more open syllables (e.g. CV) than closed syllables (e.g. CVN), so that an appearance of coda in a Japanese word is more noticeable (i.e. marked) to EJ learners. In sum, it seems reasonable to suppose that bimoraic syllables; or to be more precise, the durational contrast between bimoraic syllables and monomoraic syllables could be a strong candidate for free variables in the case of CVR and CVQ, while CVN could be a nonarbitrary choice in /N/-CV sequence, but if /N/ is followed by a non-moraic nasal or a vowel, then this part of segments are likely to become variables.

We turn now to the extendibility of Restricted-Interactive Model to L2 lexical representations. Firstly, within this model, we could account for why the novice writers' errors concerning the durational contrast of vowels appeared in our error corpus although EJ novice learners had a certain degree of moraic awareness of the CVR syllable. Recall that the variability of mora counting for CVR items is largest among the novice group, from which we could assume that this group particularly

has a variety of IL lexical representations with regard to the short and long dura-
tional contrast of vowels. Similarly, Toda's acoustically-based observations (1997)
provide evidence of variable phonological representations. In her L2 longitudinal
study, Toda illustrates a continuum of IL phonology for CVR and CVQ items such
that an 'over-exaggeration' strategy made the durational contrast gradually longer,
but it got to the point of 'overshooting' (i.e. the L2 durational contrast became
longer than that of L1), which was then modified to make it slightly shorter in the
advanced group. Secondly, if we reanalyse our EJ written errors in Chapter 3 within
the Restricted-Interactive Model, we could reduce many errors to a certain abstract
level. For example, we can more clearly present voiceless-voiced CV substitution
and vowel omission errors in 1)-a しがし ('si-ga-si' for 'si-ka-si') and 3)-b いっしょ
う ('i-s-sho-u' for 'i-s-sho') as variable representations 'si*asi' and 'i-s-sho*' respec-
tively. This chapter employs Perfetti (1992)'s notion of 'variable representations'
and considers spoken and written errors at an abstract level within the Restricted-
Interactive Model.

6.3 Research hypotheses

In Chapter 2, I touched upon how the Roman alphabet input method worked with J-
word processors, and speculated that writing the pronunciation of Japanese words
with romaji or roman characters with the QWERTY keyboard would be easier
and faster for EJ learners (see 2.2.2). This idea was based on the evidence pro-
vided by Cross-linguistic studies of L1 script effects on L2 orthography processing,
where a positive L1 transfer is likely to occur if the target language and L1 is
closer (Chikamatsu, 1996; Koda, 1989; Mori, 1998; Tamaoka & Menzel, 1994). For
example, Tamaoka and Menzel (1994) reported that alphabetic subjects had pos-
itive L1 script transfer with items written in romaji when they performed word
recognition, semantic judgments, and text reading tasks (see section 2.2.2 for more
detail). Accordingly, it is possible to assume that native orthography (romaji) is
strongly linked to phonological representations of EJ learners, and that therefore
they would make a more efficient application of their L1 knowledge such as GPC
(grapheme-phoneme conversion) rules in their L2 processing of romaji in compari-
son with hiragana. If this is the case, then we could expect an effect of positive L1
script transfer, such that a writer's romaji representation will outperform the same
writer's hiragana version of a given word. In other words, there will be fewer free

variables in the romaji representation than the hiragana representation. Theoretically, we assume that there is a single lexicon for both perception and production, but this does not necessarily imply an isomorphic representation of a given word at different levels of representations. For example, we could expect to find an instance of a quality discrepancy between phonological and orthographic representations (e.g. reversal representations). If deletion errors in written production are result of a wrong phonological representation, I predict that both spoken and written production should include equal errors. On the other hand, if deletion errors in written production are result of orthography problems, written errors in hiragana should be more frequent than spoken errors, but written errors in native orthography (romaji) should not. Thus, I hypothesized that writing a word in romaji before hiragana should reduce errors in hiragana, if the difficulty is orthography, but that if the difficulties are phonological, then writing in romaji first should make no difference to the hiragana errors.

6.4 Method

6.4.1 Participants

We tested twenty-nine undergraduate students (18 male, 11 female) at the University of Edinburgh. Most of the participants (89.7%) were drawn from the same population as those in Moraic Awareness study in the previous chapter. Thus, all of the participants in this study were native speakers of English, who largely represented two groups (Novice vs. Semi-bilingual) of English learners of Japanese at different proficiency levels: literate novice learners (N=18), and skilled advanced learners (N=11) respectively. They were given a course credit for their participation in this study.

6.4.2 Materials

Word frequency and L2 database First, I set out to compile my own frequency database by conducting word counts (i.e. raw counts) in the noun category, excluding pronouns, in the course textbooks, namely three volumes of 'ICU textbook for College Students' and the first volume of 'Basic Kanji Book'. Next, the test words were drawn from the database, and I checked: (1) pitch accent, (2) frequency (M=9.17), (3) word length in terms of the number of letters, syllables and

mora, and (4) word type whether a word consists of a long vowel (simple type) or a combination of a long vowel and other bimoraic syllables (combination type).

However, factors (1), (3) and (4) were not controlled for the following reasons. First of all, no attempt was made to control the pitch accent of each word. My reservation for not taking pitch accent into account was that it is widely known that even fluent nonnative speakers of Japanese often place the lexical accent in a wrong position, and I also noticed this tendency among our skilled participants. Secondly, the word types were not perfectly counterbalanced: the ratio between simple (N=18) and combination (N=12) types is 3:2. The main reason was to counterbalance frequency over each location. Thirdly, it was difficult to counterbalance the number of morae in the word medial position in accordance with two other positions as the target vowel in this position tends to appear in words with longer length in my frequency database. In addition to this, there are three different ways to measure word length according to the number of letters (hiragana or kanji), the number of syllables, or the number of mora. A further complication is that there is no correspondence between the number of kanji and the number of morae. For example, the 4-morae word /i.mo.o.to/ *'younger sister'* is written with one kanji, 妹, while /hi.ko.o.ki/ *'airplane'* shares the same number of morae, but it is written with three kanji, 飛行機. This means that visually longer words are not necessarily be longer than visually shorter words in terms of the number of morae. As a whole, the majority of the target items were 4-mora words (63.3%). The mean length (i.e. the number of morae) of the target words for each target position was as follows: word initial (M=3.8), word medial (M=4.8), and word final (M=3.6). Note, however, that word length seems to affect only words written in unfamiliar scripts (e.g. a katakana loanword being deliberately written by hiragana) in L1 Japanese reading (Kess & Miyamoto, 1999, p. 101), so that if the same is true of L2 reading and the stimuli are presented in a familiar script, a length effect, in theory, would not affect the outcome.

The stimuli type The production test contained thirty 3- to 6-mora Japanese words, all of which included the long vowel /o:/ in either word initial, medial, or final position (the list, see Appendix L), in which items had a mean frequency of 9.2, 9.1 and 9.3 respectively (range 1-37). The problem of how to present the target word was considered. In order to avoid the target words from giving participants a sound cue, we avoided the use of hiragana, because the direct mapping to morae would allow the participants to judge the presence or the absence of the long vowel

visually from the mora-based hiragana string. A way around this problem may be using the semantic route. In L1 studies with Japanese children, researchers normally use pictures, but it does not seem to be an ideal way for EJ participants, since some learner variability may be expected in the IL lexicon of learners of Japanese. For example, a picture of a car would recall three different synonyms - 'kuruma', 'jidoosha', or 'jikayoosha' or 'kaa' as in the English loanword. This is why I presented the target words in English translation in 4-4, yet some words also ended up with more than one version. For instance, *gold* came out with two synonyms - *kin* and *oogon*.

Another possibility for the semantic route would be to use kanji, which may provide a little sound cue at the sub-lexical level but certainly no straightforward whole word phonological information. For this reason, the targets were presented in kanji in the current study. All words were carefully selected only from course textbooks, so that the participants were expected to have representations of the target words before the experiment. Hence, no particular pre-experimental session was given for the participants to learn the target kanji items so as to avoid any practice effect. In summary, the items were comprised of two single kanji words, twenty 2-morpheme kanji words, six 3-morpheme kanji words, and 2 words that contained hiragana letter(s) due to an honorific / beautification prefix o- and a polite-ending -*san*, which are always be written in hiragana letters. To disguise the experimental items, similar items were created as fillers.

Test words and block orders The test words (N=30) consist of either one long vowel (simple type, N=18) or a long vowel and at least another bimoraic syllable (combination type, N=12). In addition, five practice materials were prepared. There were also forty fillers (Mean frequency=9.75; range 1-56) consisting of twenty CV type words, ten CVN type words, five CVV (diphthong) type words, and five CVQ type words. All fillers were also drawn from the same frequency corpus. The items were presented in two experimental blocks, each of which included an equal number of experimental items, and the order in which the two blocks were presented was counterbalanced across participants. The experimental task (see below) involved the participants writing out kanji words in two different orthographies (romaji and hiragana) after pronouncing the word, and the order of the two orthographies was experimentally manipulated as a between items factor. The orthography order was constant within any one experimental block, and changed half-way through the experiment, when the participant began the second

block. The orthography order was counterbalanced across blocks so that responses were gathered for each block both in romaji-hiragana order and in hiragana-romaji order, independently of the relative order in which the two blocks appeared. Both block-order and within-block orthography-order were balanced across subjects in the two participant groups as equally as possible. [2] A total of 73 slides (1 trial per slide, see Appendix M), including instructions, were made for each block as a series of PowerPoint presentations, in which three short breaks were inserted at equal intervals. Target words in kanji were written in 44 points, MS Mincho font.

6.4.3   Task and procedure

Participants were tested individually in a quiet room. The naming task requires the participants to name a given visual presentation of a word, and 'naming latency' or Reaction time (RT) is measured. In order to measure naming latency, a cue sound was embedded at the point where the screen showed each new target. Firstly, five practice trials were provided before the block of experimental trials to familiarize the participants with the procedures. Next, in the experimental block, the screen holds the target item for one minute, during which the participants were asked to perform a naming task (i.e. words are read in isolation), a pronunciation task (i.e. words are read in a frame sentence), and two spelling tasks. The participants were allowed to initiate a trial by hitting the enter key. Otherwise, the screen automatically changed to the following target after one minute under the control of a computer. The participants were not allowed to continue any unfinished task after one minute.

As each test word was shown on the computer screen one at a time with a cue sound, the participants were first asked to read aloud the test word in isolation as quickly and as accurately as possible. A frame sentence also appears below the target word as "kore wa to yomi-masu" (this reads as) in kana-kanji mixed text. In the following trial, this frame sentence was read with a given target word without any pause for the assessment of their pronunciation. All spoken productions were digitally recorded on SHARP MD-MT90H at 44.1 kHz sampling, and then a total of 711 samples of the test words (30 words x 29 trials, 159 no recalls) were marked for down sampling (Mono) with Cool Edit 2000. Finally, each sample was stored as a 16-bit/11.025kHz wav.-file, and subsequently the naming latency of each trial was

---

[2]In fact this counterbalancing was not totally equal because the skilled reader group included an odd number of participants.

measured from the beginning of the cue sound, at the point where low frequency range noise begins at around 1000 Hz in the wide-band spectrogram, to the onset of the word spoken in isolation.

In the spelling tasks, a booklet was prepared for each participant, in which the written order of orthography was clearly printed. The participants were asked to write down the pronunciation of the target word in both hiragana and romaji according to the order of the orthography in a given answer booklet. Once the written task was carried out, the participants were not allowed to redo the same task. For each session, I also noted any deviation from normal spelling processes. A total of 1,422 spellings (30 words x 29 trials x 2 spelling tasks, 318 blanks) were collected, which were preliminarily analyzed, coded and then electronically compiled for further statistical analysis.

6.5 Interrater reliability

This section touches upon an issue of reliability of my auditory judgments. In order to estimate interrater reliability, two native teachers of Japanese were given a random 20% sub-corpus of pronounced samples (142/711) taken from the entire spoken samples, and asked to judge the length of every vowel in a given word as correct, shortened or lengthened. Thus, their focus were not on a particular item, namely the long vowel /o:/ in their auditory judgment task. The agreed rate, including judgments of non-target vowels, among three raters was very high 92.3% (131/142). A further statistical test was carried out for the estimation of interrater reliability on the target vowel. This was measured by Kappa coefficients, which estimates whether interrater agreement is beyond chance for binary and categorical ratings. The reliability Kappa scores of my judgment with two raters were .74 and .72 respectively, indicating very good agreement (Fleiss, 1981).

6.6 Results and discussion

6.6.1 The pronunciation task

Of 870 trials (30 targets x 29 trials), 63% (N=548) of the total pronunciation task was identified as correct word pronunciation in my auditory judgment. Table 6.1 below shows an outline of the results of the pronunciation task by each subject

group. Overall, the correct percentage of the skilled learner group reveals that this group recalled more words, showing an ability to name words more accurately than the novice group.

Table 6.1: Outline of the pronunciation task by each subject group

	Novice (N=540)	Skilled (N=330)	Collapsed (N=870)
No recall	28.7 % (155)	1.2 % (4)	18.3 % (159)
Errors	23.5 % (127)	10.9 % (36)	18.7 % (163)
Correct	47.8 % (258)	87.9 % (290)	63.0 % (548)

In Table 6.1, 'Errors' include both 'completely wrong' and 'partially wrong' errors (i.e. a target-like pronunciation consisting of the same number of mora or half of the kanji compound was wrongly pronounced). In the novice group, 'Partially wrong' items comprised 78 out of 127 errors. In other words, 37.8% of the pronunciation sample (i.e. a sum of 'No recall' and 'Completely wrong'=204) in the novice group corresponded to a lack of lexical entries for the given set of words.

Next, Table 6.2 shows a summary of target vowel omission[3] rates for each subject group in the pronunciation task.

Table 6.2: Summary of target vowel omission rate

	Novice (N=540)	Skilled (N=330)	Collapsed (N=870)
Correct	54.3% (293)	90.9% (300)	68.2% (593)
Omission	7.0% (38)	4.2% (14)	6.0% (52)
Other errors	38.7% (209)	4.9% (16)	25.9% (225)

Again skilled learner group showed greater accuracy on the durational control of the target vowel. As for the novice group, it is interesting to compare 'Correct' target vowel (N=293) in Table 6.2 with the 'Correct' word pronunciation (N=258) in Table 6.1. This discrepancy may indicate some other problems rather than the durational control of the target vowel itself. We shall come back to this point later.

---

[3]Throughout this chapter, I use the term 'Omission' or 'target vowel omission' to refer to the omission of the target (i.e. the second element of a VV sequence). Thus, it should mean that the long vowel (VV sequence) is shortened rather than dropped altogether.

6.6.2 The naming task

In the naming task, the participants of two different proficiency groups read the same experimental material, which produced total of 711 tokens in 870 trials. Then the naming latency or Reaction times, which includes time for lexical access and vocalization, of a given word of each token was measured. As is often the case, the naming task in the current study involved identifying[4] words in isolation in order to avoid confounding variables that are associated with sentence context.

In general, mean reaction times of novice readers (8,021.75 ms; SD=8,322.36) were more than two times slower than those of skilled readers (3,622.52 ms; SD=2,693.64). This difference was significant by subjects and items ($t1(27) = 4.25$, $p < .001$, $t2(29)=6.80$, $p < .001$). This difference indicates that skilled readers recognized given words much faster than novice readers, as well as taking less time to read them aloud after identifying them. A Pearson's correlation was performed to test whether there was any relationship between reaction times and word frequency. For this analysis, incorrect responses were removed. There was a significant negative correlation for novice readers group ($r= -0.62$, $p < .05$), while we found a marginal negative correlation ($r=-0.31$, $p =.09$) for skilled readers group.

Figure 6.1: Scatter plot of latency x frequency
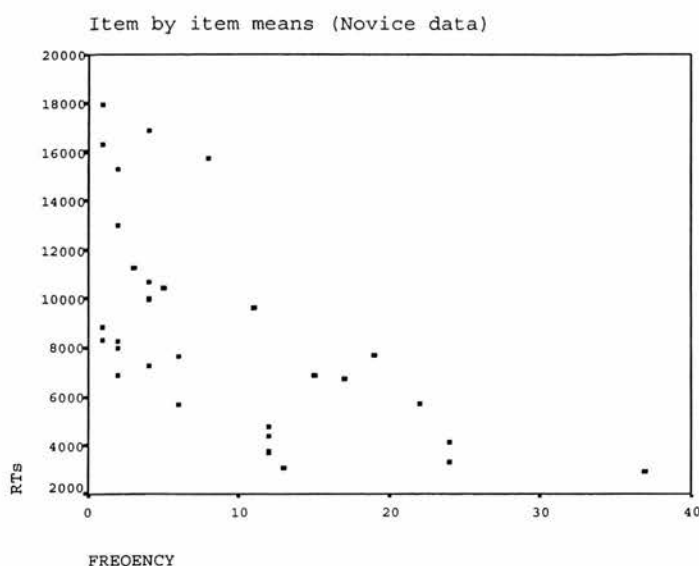
Moreover, the correlation between the rate of correct pronunciations and word frequency was positive and significant for the novice group ($r=0.51$, $p < .05$), but was

---

[4]I use 'recognition' and 'identification' reciprocally in this study.

not significant for the skilled readers group (r=0.17, p =.37). This demonstrates that the more experience they had of the words, access speed and reading accuracy increased in both groups, though the strength of the effect is obviously different between novice and skilled readers. In other words, we found evidence for 'frequency effect', that was greater for the novice group. However, we may need to consider if there is any word that shows sign of a speed-error trade-off, which is an instance of increase in error rate rather than slower the reading speed, in the data. Some words will be considered and discussed later.

Table 6.3 reports mean reaction times (ms) and percentages of errors in each group, showing data for the words which were pronounced with the highest and lowest accuracy. The rank was made according to the correct percentage of word pronunciation collapsed over both groups. Table 6.3 largely demonstrates that the participants took a longer time to identify difficult (i.e. higher errors) items than easier (i.e. fewer errors) items.

Table 6.3: Mean RTs (ms) and percentages of errors in each group

	Word frequency	Novice RT(error)	Skilled RT (error)
1. 'nitiyoobi' (*Sunday*)	24	4,119.78 (0)	2,838.73 (9.1)
2. 'benkyoo' (*study*)	37	2,918.63 (16.7)	2,779.64 (0)
2. 'gakkoo' (*school*)	12	4,347.67 (16.7)	2,551.64 (0)
28. 'shoosetu' (*novel*)	2	7,968.80 (88.9)	3,527.73 (18.2)
29. 'koosoku' (*high speed*)	1	16,307.38 (94.4)	3,673.67 (45.5)
30. 'isshoo' (*one's life*)	1	8,839.06 (88.9)	3,894.36 (63.6)

For example, the word 'koosoku' (*high speed*) exhibited long reaction times among novice readers, and high error rates. There was only one novice reader who correctly pronounced this word, obtaining the shortest response time (2,000 ms) in this group, while remaining readers took a much longer time to name the word. This slow response with higher error rates among novice readers seems to be caused by the low frequency of 'koosoku' (*high speed*). However, another equally infrequent word ('isshoo'), showed a much shorter latency (8,839.06 ms) compared with 'koosoku' (16,307.38 ms). This trend was seen not only for the novice data, but in the means for all subjects ('isshoo' (6,896.50 ms) and 'koosoku' (9,618.94 ms)). Figure 6.2 plots mean accuracy (percentage correct pronunciation) against mean reaction time for each item, collapsed across all subjects.

Figure 6.2: Percentage correct pronunciation
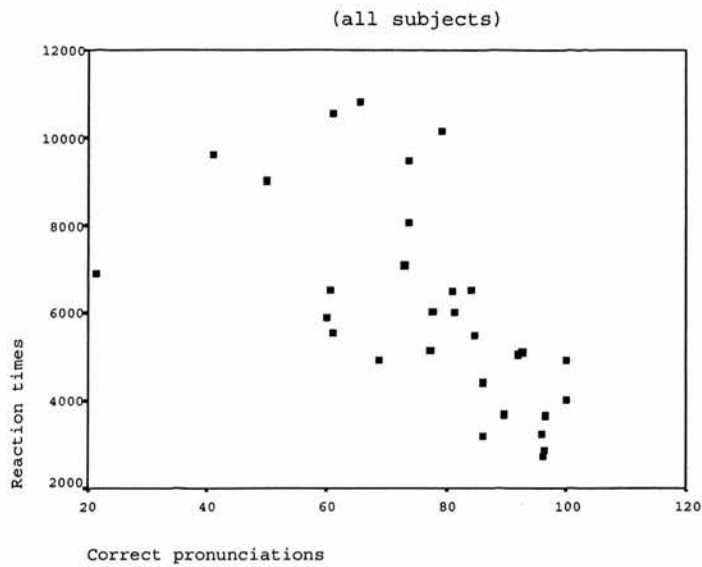
The figure 6.2 shows that reaction time correlated negatively with accuracy ($r = -.62$, $p < .05$). However, it is clear that the word 'isshoo' represents an outlier in the scatterplot, with a mean accuracy of only 21.4%, but a mean reaction time of only 6897 milliseconds (the relevant data point can be seen in the extreme left portion of the graph). Removing this outlier resulted in a negative correlation of a greater magnitude ($r = -.71$). This suggests that the word 'issho' led to a speed-accuracy tradeoff, and we will examine this possibility in more detail in the following paragraphs.

Consider Table 6.4, which lists every word from the current study with frequency-1, along with a break-down of error types. Mean latency of each word was as follows: 'koosoku' (Novice:16,307.38; Skilled: 3,673.67) 'ginkooin' (Novice: 8307.07; Skilled: 5428.45), 'tashoo' (Novice: 17950.75; Skilled: 4646.40), and 'isshoo', shown in Table 6.3 above.

Table 6.4: List of word frequency-1 items

	Correct	Wrong	Partially wrong	No recall
'koosoku' (*high speed*)	1 (6)	3 (0)	4 (3)	10 (2)
'ginkooin' (*bank clerk*)	8 (11)	0 (0)	7 (0)	3 (0)
'tashoo' (*more or less*)	2 (9)	4 (0)	2 (1)	10 (1)
'isshoo' (*one's life*)	2 (4)	3 (0)	12 (7)	1 (0)

(Note. The numbers given in parentheses are the results of skilled readers)

Table 6.4 shows that the number of 'Partially wrong' responses was higher for 'isshoo' than the other words of the same frequency ($\chi^2 = 16$, p <.01, combining counts from both groups of participants). All errors in this category were read as 'issee' by both groups. In Japanese, the kanji morpheme 生 in 一生 /issjo:/ has a reading that alternates between /sjo:/ and /se:/ (for example, 生産 'see-san' (*production*) and 生涯 'shoo-gai' (*lifetime*)).

Although EJ readers in the current study were taught about a pair of alternative readings for this kanji, hence, had linguistic knowledge, the mean latencies suggest that readers did not perform a lexical check to see if there were any alternative pronunciation codes. This could be seen as a result of speed-accuracy trade-off; or, to put it another way, the readers were fairly confident about the pronunciation, so that they identified the word quickly as 'issee', but as a result of their belief, they made more errors than they did for other words of the same word frequency. A potentially important factor related to this is the distribution of the /se:/-/sjo:/ alternatives in both learners' inputs, namely in our course text frequency database for novice readers, and natural language data for skilled readers. First of all, in our frequency database, words that include 生 use the /se:/ pronunciation of this character 162 times, while the /sjo:/ pronunciation is only counted 4 times, showing a very skewed token frequency distribution. Type frequency is similarly skewed: my word count in a kanji dictionary (Oobunsha Kanwa Jiten) points to the same tendency, in that 85.1% of given examples in the dictionary use the /se:/-pronunciation (57/67), while the words that use the /sjo:/-pronunciation (6/67), or that allow both alternatives (4/67) are very few. Thus, EJ readers performance seems to accurately reflect the skewed occurrence of /se:/-/sjo:/ alternatives in their course textbooks as well as type frequency in natural language data. In short, this example provides evidence for how an exception word creates L2 reading difficulty, which may not be associated with particularly high latency.

6.6.3   The spelling task

One of the questions addressed by this study is whether or not writing a word in romaji before hiragana leads to reduce errors in hiragana productions. To answer this question, we collected a total of 1,422 spellings (30 words x 29 trials x 2 spelling tasks, 318 blanks) from 29 English-speaking learners of Japanese, and statistical analyses were conducted to assess the effect of writing order. Figure 6.3 below shows target vowel accuracy for romaji and hiragana in each writing order.

Figure 6.3: Target accuracy by writing order for romaji (R) & hiragana (H)

Two-way analyses of variance were carried out with the variables of learning group (i.e. novice vs. skilled writers) and writing orders (romaji-hiragana vs. hiragana-romaji) on the dependent variable of percentage of target vowel omission in either the romaji or the hiragana orthography. Both participants (F1) and items (F2) analyses were calculated. The first ANOVA, analyzing omission rates in the romaji data confirmed that writing order was not significant, but there was a main effect of subject group, with a lower rate of target vowel omission for skilled writers than novice writers, though this was marginal in the analysis by subjects (F1 (1,27) = 4.12, Mse = 216, p = .052; F2 (1,29) = 8.45, Mse=177, p <.01). The second ANOVA with hiragana data also confirmed the same trend (F1 (1,27) = 2.63, Mse = 268, p = .12; F2 (1,29) = 7.876, Mse=144, p <.01). The effects of order and interaction between order and subject group were not significant for either romaji or hiragana data (all F's <1). Thus, it was confirmed that writing order has no effect on omission rates for either novice or skilled writers, though omission rates did differ between the groups. Fuller discussion will be presented in the following section.

Next, we shall consider how spoken and written productions coincide with each other. Table 6.5 below reports percentage of agreement in spoken production (P), romaji spelling (R), and hiragana spelling (H) by two writer groups. The highest agreement in both groups is the 'PRH agreed' category where pronunciation, romaji spelling and hiragana spelling all agreed in both directions. In this category, if pronunciation was correct, spelling was also correct for a given word. Likewise, if

pronunciation was incorrect, spelling coincided with the spoken error. This means that regardless of the romaji or hiragana spelling of a given word, the written production of participants from both groups very often matched their pronunciation.

Table 6.5: Percentage of agreement in spoken production, and romaji and hiragana spelling

	Novice (raw number)		Skilled (raw number)	
P only	1.5	(8)	2.4	(8)
PR agreed	2.2	(12)	1.5	(5)
PH agreed	1.7	(9)	1.2	(4)
RH agreed	3.7	(20)	2.4	(8)
PRH agreed	89.6	(484)	91.5	(302)
PRH disagreed	1.3	(7)	0.9	(3)

In section 6.3, we predicted that both spoken and written production should include equal errors, if deletion errors in written production are the result of wrong phonological representation. Thus, by definition, a large proportion of written production (89.6%: novice group and 91.5%: skilled group) reflects the writer's phonological representation, suggesting that phonological and orthographic representations are isomorphic. However, spoken and written production did not include an equal number of target vowel omission (i.e. shortened) errors - the omission rate in spelling (N=52, N=22) outnumbered the omission rate in pronunciation (N=38, N=14) for novice and skilled writers respectively. In addition, both groups included additional illegal hiragana errors (Novice: N=5; Skilled: N=6), in which the target vowel was present in both spoken and written productions, but a wrong hiragana letter was used for the target vowel. These results too coincide with another prediction, namely that written errors in hiragana should be more frequent than spoken errors, if deletion errors in written production are result of an orthography problem.

Though the number is small, more frequent hiragana errors confirm that regardless of proficiency level, EJ writers have a persistent orthography problem for some words. For example, in Table 6.3 above, 'gakkoo' (*school*) was ranked as the second best pronounced word. However, this goes two ranks below in romaji representation, and then drops further to the ninth in the hiragana rank. A typical spelling error was a missing final vowel letter as in 'gakko*', but there were some other errors including a missing small hiragana letter っ at the first half of geminate consonant position as 'ga*koo' and its voiceless version '*a*koo' by missing two voicing dots

for the first hiragana letter. As has been mentioned above, the novice group's spoken errors include a fairly good number of 'partially wrong' errors (61.4% of total errors). Together with a high PRH agreement rate (89.6% of the total) we can be pretty certain that novice learners' orthography representations contain many free variables besides the target vowel omission. On the contrary, the high accuracy of skilled learners' word pronunciation (87.9%) and a rather high PRH agreement (91.5%) strongly suggest that skilled learners have acquired fully specified phonology and orthography representations of a given set of words in this study.

6.6.4   Lexical acquisition order

This section examines lexical acquisition order of the Japanese words that were used in the reading experiment. An implicational scale (see Figure 6.4 on the next page) was made that allows us to locate each participant's data and degree of difficulty for each individual item in one distribution (Hatch & Lazaraton, 1991). This technique involves the construction of a matrix in which every participant's results are rank-ordered, so that participants in the matrix appear with the best participant at the top and lowest at the bottom. In addition, each item is rank-ordered along the horizontal axis from the most difficult (left) to the easiest (right). The standard use of this method requires dichotomous (yes/no) sorting of raw data of the participants' responses into 1: 'yes' (Correct: 'acquired'), and 0: 'No' (Wrong or No recall: 'not acquired'). A major problem in using this technique is that the mandatory dichotomous sorting will result in the loss of most interesting 'Partially wrong' or variable representations in our data. In order to accommodate this part of the data without violating the procedure itself, I abandoned the idea of plotting the data with figures 1 and 0; instead I used a colour scheme - (1) all 1-items were presented in white (i.e. acquired), (2) all 'Wrong' or 'No recall' (i.e. not acquired) data in 0-category were displayed in black, and (3) all 'Partially wrong' (i.e. incomplete IL representations) data in 0-category were expressed in gray. Thus, we did not conduct the usual final procedure of drawing a clear-cut line between acquired and not-acquired items. To summarize, the modified implicational scale was based on dichotomous data, but the 0-catogory was manually divided into two levels in the final plotting, so that the modified matrix shows white, black and gray areas representing the acquisition status of the items.

Figure 6.4: Ranking of participants (vertical) by items (horizontal)
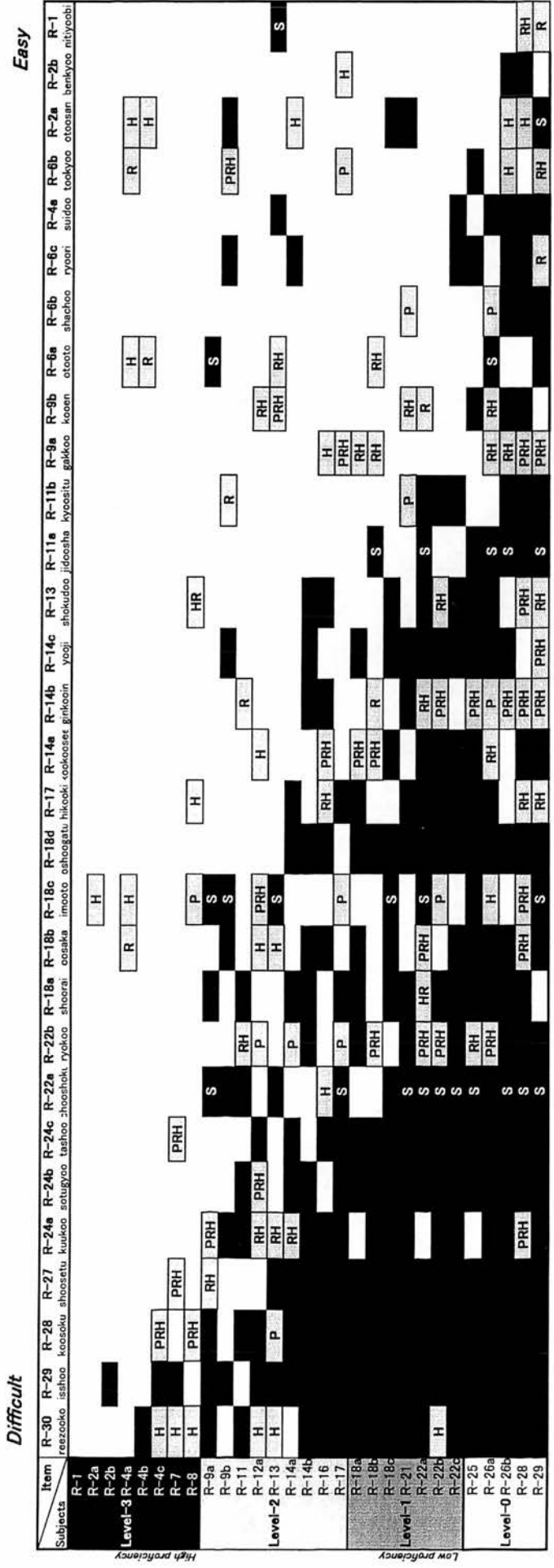
Note: black = No recall, gray = Partially wrong, and white = Correct. S = semantically related error, P = pronunciation, R = romaji, H = hiragana (e.g. PRH means all three are partially wrong).

In this scaling, I have also shifted my point of view from dichotomous grouping of Novice and Skilled learners groups. As we can see from the scaling, participants were ranked according to their all PRH correct data from highest (100%) to the lowest (10%). Levels are represented as Level-0 (the lowest proficiency group) to Level-3 (the highest proficiency group) according to their ranking. Although the majority of participants stayed largely in the same proficiency range, the 4 levels presented in the implicational scale provide us with more information about the participants' relative skills. For example, an exceptionally well-performing novice participant ranked 5th, and another novice participant ranked 11th in the scaling, while a reader from the skilled group appears on the borderline between Levels 1 and 2. To summarize, participants in the Level-0 range identified very few words properly (M=19.33, SD=7.23), while participants in the Level-1 range performed better (M=40.48, SD=6.79). In Level-2 (M=60.74, SD=7.22) participants who achieved 50-60% or more words were located in the lower range (M=54.17, SD=3.19), and between 60-70% scorers were in the upper range of Level-2 (M=66, SD=4.35). Level-3 (M=91.26, SD=6.41) participants identified 80% or more words.

A loss of mora

Appendix-N shows the ranking of all the PRH correct words that appear horizontally in the implicational scale. I have added extra information (kanji and romaji[5] with word meaning) to the ranking below. An asterisk on romaji indicates irregular hiragana usage, and word frequency of each word is given in parentheses.) Interestingly, words at the top ('nitiyoobi') and the bottom ('reezooko') in the ranking are 5-mora words with three kanji morphemes. There are two crucial differences between the two words. Firstly there is clearly a difference in their word frequency. Secondly, their word type is different - the former contains only one long vowel /o:/ (simple type), while the latter contains two long vowels /e:/ and /o:/ (combination type). There was an instance of reduction of mora in 'reezooko' (*refrigerator*) from 5 to 4, resulting in the form re-e-zo-ku (i.e. re-e-zo*k*), in which the duration of the word initial vowel /e:/ was successfully controlled, but the word medial vowel /o:/ was shortened, leading to the loss of a mora, and the vowel in the word final CV mora was also underspecified. Two Level-2 readers (R-9a and R-14b) made this complex reduction error, and their spellings were also identical to their spoken production. As this example illustrates, combination type words are generally considered to be more difficult than simple type words. Nevertheless, the correct

---

[5]Geminated sounds are written as doubling the same sound. Thus, they do not follow hiragana convention unlike roman input of J-word processing.

percentage of each type was not very different in this study (Simple: 56%, Combination: 56.4%). The difference was that combination words tended to attract more partial errors (15.9%) than simple words (9.8%). In our perceptual study in the previous chapter, the novice group also seemed to be suffered from perceptual overload from the increased number of syllables in the devoiced condition. Likewise, Chikamatsu (1996) reported an effect of word length in kana word recognition, such that English readers slowed down as the number of kana letters increased. It is not necessarily for a combination word to be longer than a simple word in length, but the complexity of words with longer length could be an important factor determining the quality of IL lexical representations of some words.

Hit-or-miss items and lexical decomposition strategy

In English, a learner's vocabulary growth implies an increase of phoneme-sound mapping rules. Experienced English readers may correctly compute unfamiliar or irregular words directly from spelling by applying more accurate and powerful GPC (grapheme-phoneme conversion) rules. In contrast, reading kanji compound words in Japanese is, in essence, a hit-or-miss affair (see 2.1.4), at any level of L1 and L2 learning. In a nutshell, the decoding power of 'reading for meaning' increases in proportion to the number of acquired kanji words, but as far as 'reading aloud' (i.e. phonological coding for pronunciation) is concerned, we do not have backup mapping rules to compute word phonology from a given kanji compound.

What if a situation forces a reader to read a never-encountered kanji compound word? Though such occasions are fewer in the case of L1 adult reading, native speakers can often roughly estimate the most likely match from sub-lexical level information from each constituent kanji in the compound. Thus, low frequency Japanese words are likely to be read in a hit-or-miss manner by means of unreliable sub-lexical matching strategies. In the current study, for example, we have already mentioned the slow response and high error rate for the low frequency word 'koosoku' (*high speed*) at R-28. All underspecified errors 'koo*oku' in gray cells at Level-3 were very minor segmental mistakes, where a voiced consonant was given instead of a voiceless one (/ko:zoku/ for /ko:soku/). On the other hand, all lower proficiency readers appeared to employ a sub-lexical decomposition strategy in their attempt to access the phonological code of *high speed*. Reading errors suggest that they broke the compound kanji 高速 /ko-o.so-ku/ into two components 高 and 速. Each kanji has both an on-reading and a kun-reading: they are 高 (/ko:/ or /taka/)

and 速 (/soku/ or /haya/) respectively. Thus, errors of the first kanji naturally included both /ko:/ and /taka/. Note, however, that the second kanji activated more variables than the first component. These were, (1) /osoi/, (2) /sju:/, (3) /do:/, and (4) completely irrelevant sounds such as /isu/. The error in (1) is a straightforward case of antonym activation, in which the kun-reading of a kanji with the opposite meaning 遅 /oso(i)/ (slow) was named. The antonym shares the left-hand side (i.e. the left-bottom radical called 'nyoo') of the second kanji of the target compound 速. Similarly, if we compare the left-hand side of 速 with another possible kanji [6] 週 (/sju:/ week) and 道 (/do:/ *road*) for the above phonological errors, we notice that both kanji also share the same radical of the second kanji 速. This could be an instance where a sub-component of the second kanji activated the phonological code of a word that shares the same left-bottom radical. If this is the case, then Level-1 and some lower Level-2 EJ readers broke up the kanji compound 高速 twice; once at lexical level, and once at the morpheme level. Thus, the three errors provide evidence for lower proficient EJ reader's decomposition reading strategies, and they also suggest a possible perceptual grouping mechanism for reading single kanji, and the salient features for early L2 kanji recognition.

Strength of visual-phonological mappings and graphic interference

Next item to consider is 'imooto' (*younger sister*), which is represented by a single kanji 妹. This kanji is always introduced in a very early stage of our first year course together with other family terms such as brother, father, and mother. We counted the word 15 times in the course textbooks. A similar item in the list is 弟 'otooto' (*younger brother*) that has a slightly lower word frequency count of 12. The two words were introduced at the same time and word familiarity, word frequency, and the word lengths are almost equal. However, 'imooto' (*younger sister*) was ranked at 18th, while 'otooto' (*younger brother*) was ranked at 6th in the ranking. The only difference between the two family terms seems to be their degree of confusability with semantically similar kanji. The kanji 弟 'otooto' (*younger brother*) is visually highly distinct from 兄 'ani' (*elder brother*; frequency=12), and the two kanji do not share any sub-components. On the contrary, the sister-pair 妹 'imooto' (*younger sister*) and 姉 'ane' (*elder sister*; frequency=10) share the left-hand radical 女 (*woman*), while differing only in the right-hand component 未 or 市.

---

[6]/sju:/ hit 42 homophonic kanji when write with a J-word processor. Similarly, /do:/ hit 19 homophonic kanji. However, 週 /sju:/ and 道 /do:/ are the only kanji that consists of the left-bottom radical identical to the target 速 /soku/ among those homophones.

Evidently, throughout all proficiency levels except Level-3, 'imooto' (*younger sister*) activated and resulted in the recognition of a semantically and visually close 'ane' (*elder sister*) four times, and twice resulted in the recognition of the semantically similar, but visually distant 'ani' (*elder brother*). On the other hand, there was only one instance of 'otooto' (*younger brother*) being identified as 'ani' (*elder brother*) by a Level-2 (R-9a) reader, and an instance of it being identified as 'ane' (*elder sister*) by a Level-0 (R-26b) reader. This difference gives us the idea that the brother-pair has more stable connections of visual-phonological mapping than the sister-pair. Moreover, against traditional belief that assumes a single kanji pair such as the sister-pair is easy to learn, only among Level-3 did readers hold stronger connections of visual-phonological mapping in lexical representations of the sister-pair. Thus, seemingly easy 1-morpheme kanji was found to be a very confusing lexeme in the semantic field of 'family members'. We also found similar evidence for semantically related wrong identifications (see below).

Similarly, 'chooshoku' (*breakfast*) at rank 22 includes many synonym activation errors, which are marked as 'S' in the implicational scaling. Table 6.6 reports a summary of words that share the same word frequency of 4 with 'chooshoku'. Table 6.6 shows that 'chooshoku' includes numerically more errors than the other items. It is important to note that there was only one instance of an underspecified phonological representation (/sju:sjoku/) in 'chooshoku' (*breakfast*) trials in the novice group, unlike other words in this study.

Table 6.6: Summary of words that share the same word frequency of 4

Rank	Item	Mean RTs (mse)	Correct	Wrong	Partially wrong	No recall
6	'shachoo' (*president*)	6,002.81	75.9	3.4	13.8	6.9
18	'oshoogatu' (*New year*)	9,503.32	48.3	3.4	13.8	34.5
22	'chooshoku' (*breakfast*)	9,005.75	48.3	27.6	20.7	3.4
24	'sotugyoo' (*graduation*)	5,536.50	37.9	3.4	17.2	41.4

(Note. The numbers except mean RTs are percentage of each item)

The remaining errors (N=10) were all related either to the semantically and phonologically similar 'chuushoku' (*lunch*), or to a synonym of *breakfast* in kun-reading (e.g. 朝ご飯 asa-gohan; frequency=9), all of which were pronounced accurately. Although both 'chooshoku' (*on*-reading) and 'asagohan' (*kun*-reading) are alternatively used both in spoken and print, the *on*-reading is slightly more formal and often appears in print, while the *kun*-reading is a typical spoken feature. Interestingly, three skilled readers in Level-2, whose print exposure is much greater than

that of the lower proficiency readers, also made exactly the same kind of errors - one as an underspecified phonological representation /sjo:sjoku/, one for kun-reading, and one for semantically related 'yuushoku' (*supper*).

In Chapter 3 (Error Analysis), we found some mysterious kanji errors, from which we cannot trace the meaning of the intended words even with the help of sentence context. We have identified two potential candidates for such kanji in the 'chooshoku' (*breakfast*) trials. The first one is 就職 /sju:sjoku/ (*finding employment*) and the second one is 少食 /sjo:sjoku/ (*light eating*) if underlying imprecise phonological codes are used as an input for the intended word /tjo:sjoku/ (*breakfast*) in J-word processing. At first glance, the errors /sju:sjoku/ and /sjo:sjoku/ do not indicate much difference, but according to the above analysis, the former is an underspecified representation of a semantically related 'chuushoku' (*lunch*), whilst the latter is an deficient representation of the target word 'chooshoku' (*breakfast*). Under this view, a Level-2 error of /sjo:sjoku/ reflects a better variable representation (i.e.*ho-o-sho-ku) than a Level-0 error of /sju:sjoku/ (i.e.*h*-*-sho-ku) for the intended word. However, another possibility would be to argue for 'imprecision' in terms of a lack of articulatory control of the onset affricate, rather than deficit in their lexical representation. For example, in /sjo:sjoku/ it could be seen that the wrong manner of articulation was employed, in such a way that the intended palato-alveolar affricate became a palato-alveolar fricative at the same palatal region of articulation. Nevertheless, this is unlikely to be the case, as the romaji and hiragana production also used the representations identical to the spoken error. Thus, we may conclude that this example has demonstrated that even phonologically very close errors reflect different developmental stage of variable representations.

Reversal representation and an overshooting strategy

Finally, we shall consider some instances of what Perfetti (1992) calls "reversal" representations, in which the phonological representation is fully specified, but this representation is represented orthographically by an illegal letter. We will explain how such reversal representations develop from less proficient to more skilled learners, using examples from the experimental data. The first kind of error, which is common in very low proficiency learners, is exemplified by a level-0 reader, who identified 'imooto' without the target vowel as /imoto/, then spelled it as i-mo-to in both romaji and hiragana (i.e. いもと). This involves the representation 'imo*to'

caused by the loss of a mora (a missing letter). In slightly more proficient learn-ers, one finds errors which are exemplified in our data by an instance at Level-1, where the phonologically accurate /imo:to/ was correctly spelled in romaji (i-mo-o-to), but where in hiragana it was written as いまうと (i-ma-u-to). Here, 'im*oto' included the target vowel, but the error was caused by a variable in the second mora. Finally, there are reversal errors, in very proficient learners. These errors are exemplified by a hiragana error いもおと (i-mo-o-to) that was made by two very proficient writers at Level-3, who were ranked at R-2 and R-4. This error involved the use of the alternative お instead of the default う for the second element of /o:/. This representation 'imo*to' has a precise phonological representation with a wrong letter for the target vowel. In short, the first and the third types of errors share the same representation. Thus we found evidence of a reversal representation.

A similar reversal error was found in /oosaka/ trials. This time the spelling error involved use of the default う instead of the correct alternative お (i.e. an exceptional word). Learners have explicit knowledge that most of the time the default is the う-option, but a complication is that this spelling ambiguity cannot be solved by any learnable rule. In the present study, the error pattern of a proficient writer ranked as R-4b demonstrates his writing strategy. This writer identified target words 100% correctly, but made hiragana errors in the easy-learning range. His problem was overshooting. This proficient writer employed less-used お-option whenever he felt uncertain about which one of hiragana was the right version for a given word. His errors include いも お と (i-mo- o-to: *younger sister*), おと お と (o-to- o-to: *younger brother*), and おと お さん (o-to-o-san: *father*), all of which should have been written with the う-option. However, the same writer spelled 'oosaka' (*Osaka*) without any mistake as おおさか, and this was the only word in the current study to require the お-option. This pattern of error suggests that he was too aware of the う-お alternation, and consequently made these overshooting errors. As I have pointed out, teachers of Japanese seem to regard this spelling difficulty as an early learners' problem, but the present study found not only that inexperienced lower proficient writers were affected, but also that some skilled writers too suffered from spelling ambiguity between the two options, even well after acquiring fully specified phonological representations.

6.7 General discussion

In this study, we investigated EJ learners' interlanguage lexical representations in reading and spelling. We posed the general question of whether writer's linguistic problems are reflected in their written products, and we were especially concerned with the nature of the relationship between phonological and orthographical representations in terms of the loss of a mora in the long vowel /o:/. The results of the pronunciation task revealed that participants in the skilled readers group identified nearly 90% words correctly, indicating the number of fully specified phonological representations of a given set of words. On the contrary, 37.8% of novice samples were found to have no lexical entry, and about 60% of total errors included partial errors that suggest the existence of variable IL phonological representations among the novice readers group.

The reading tasks

The results from the pronunciation task and reading tasks produced two key findings to answer our questions. First of all, we found evidence that spelling and reading share the same lexical representation. There are, however, apparent individual differences in our L2 data, such that the phonology-orthography agreement rate did not increase according to the proficiency continuum. Some skilled readers still had a graphic representation deficit despite their having fully specified phonological counterparts. On the other hand, even some of the low proficiency readers developed a higher rate of perfectly agreed phonology-orthography lexical representation though the number of acquired items was very small. Thus, we confirmed largely two types of L2 lexical development patterns.

The first type is a phonology-then-orthography acquisition. This is a typical L1 development pattern. The second category is a synchronized phonology-orthography acquisition type. This is unthinkable for L1 learning, but it is a typical L2 lexical learning, as spoken and written forms of a given word are likely to be taught simultaneously. It is also possible for an IL lexicon to include words belonging both to type 1 and 2, since vocabulary learning takes place basically as a word-by-word process. Likewise, our hypothetical syllabeme IL language in 6.2.1 is not evident according to the learner proficiency group, but rather all lexical representations seem to go through a 'syllabeme stage' (i.e. syllable-based IL lexical representation) at some point in the developmental process, but its progress to reach next 'mora stage'

(i.e. a more fully specified mora sensitive IL lexical representation) may be affected by the internal structure and the length of each word.

To determine the exact nature of the quality of lexical representation of a given word, I used a naming latency technique. There is a fair body of research, which agrees that lexical access occurs within the lexicon, and that word recognition is not influenced by information outside lexicon (see for instance Seidenberg et al. (1984)). In this way, we recognize words rapidly without having interference form higher-levels of processing such as syntax, semantic and pragmatics. However, Perfetti made a point that 'interactive and autonomous processes coexist' (Perfetti (1992, p.149)). This means that while outside information cannot easily break into the lexicon during word recognition, within the lexicon it allows the inside information to be triggered interactively. For much of the skilled readers' data, we can assume that word recognition was highly automatic, so that the findings of previous research should be relevant to the interpretation of this part of our data. However, we need to be aware that more than half of the data from the novice group would be irrelevant to the above view. We shall come back to this point later.

The following is a brief summary of findings in the present study. EJ learners found the left conditions easier than the right conditions when they read and wrote:

Word frequency	High >Low
Reading options	Regular >Exceptional
Internal structure	Simple >Complex (phonologically or morphologically)
Orthography	One-to-one mapping >One-to-two mapping

In the naming task, I found two types of evidence that the well-documented frequency effect played a role in our L2 Japanese word recognition. First, a significant negative correlation confirmed that EJ readers' reading times decreased as word frequency increased, and vice versa. Similarly, we found evidence in the rate of correct pronunciation was correlated positively with word frequency in the novice group. The more they had seen the words, the more accurate their readings became; in other words, we found evidence that word frequency affects the quality of early IL lexical representation. Note that this frequency correlation was more robust for the novice group. Skilled readers have greater exposure to authentic printed material that goes beyond course textbooks, thus the textbook-based word count had no

effect on them. The skilled readers read targets much faster[7] and more accurately. Second, regardless of their reading proficiency, EJ learners made more errors when they tried to read exceptional words, and their mistakes were found to be a reflection of a skewed distribution of reading options. From a pedagogical point of view, material writers can prepare input in terms of word frequency in a controlled manner, so that we can expect facilitative effects as well as predicting acquisition order according to a planed input. Similarly, teachers need to pay special attention to a low frequency word that has reading options with a biased distribution of each alternative.

Researchers working on spoken and word recognition (e.g. Shelton and Martin (1992) take 'semantic priming' effects into account on a par with the word frequency effect. The facilitation effect occurs if the semantically related first word (i.e. prime) speeds up the identification of the second word (i.e. target). On the other hand, if the prime slows down the reading speed of the target, the effect is called 'inhibition'. Other research such as Shelton and Martin (1992) found that semantic priming is limited to a pair of words that are linked by semantic associations (i.e. associative semantic priming), not by the words that share the word meaning. For example, the 'doctor-nurse' pair is often cited as a typical associated word pair. Many of us naturally name 'nurse' when we are asked to say an associative word of 'doctor'. Such word association effects are considered to occur via frequent co-occurrence of words rather than via a semantic network. The above claim is that such associative word pairs participate in priming effects, while semantically related non-associative words such as the same super-ordinate lexemes or hyponymy (e.g. 'car' is a hyponym of 'vehicle'; 'car' and 'bicycle' are co-hyponyms) cause no priming effects beyond any co-occurrence associations that they might have. Note, however, that this applies only to automatic word recognition processes. To discuss our L2 data, especially the novice data, this point is important. As has been shown, the novice data includes rather slow, probably non-automatic, word identification of many underspecified representations. For the Restricted-Interactive model, such representations are stored in a functional lexicon where higher-level information such as 'knowledge', and 'expectations' is penetrable to the lexicon in the word identification process. Thus, if semantic priming occurs during such non-automated slow identification, we would expect the involvement of higher-level operations outside

---

[7] We lack comparative L2 data for comparison; hence, I cannot say for certain whether participants in the current study generally read the target words faster or slower than other L2 learners of Japanese.

the lexicon such as readers' expectations or verification, regardless of whether it is associative or non-associative semantic priming.

As discussed in 6.6.4, the current study provided several lines of evidence that both phonological and semantic information played a role in the process of L2 word recognition. Especially, we found more errors that were made as a result of inappropriate lexical access between a pair of words (e.g. the sister-pair) where phonological/visual and semantic relatedness was evident. Beside the sister-pair, other co-hyponyms such as the *father-mother* pair of 'parents' ('otoosan' お父さん vs. 'okaasan' お母さん), the *president-section chief* pair of 'job titles' ('shachoo' 社長 vs. 'kachoo' 課長), the *high school-university students* pair of 'students' ('kookoosee' 高校生 vs. 'daigakusee' 大学生), and the *car-bicycle* pair of 'vehicle' ('jidoosha' 自動車 vs. 'jitensha' 自転車) were activated, and ended up with wrong identification (i.e. the right hand side of the pair).

All of the above co-hyponym pairs include one to three common kanji morphemes or hiragana, and relevant phonological information. Although these co-hyponym pairs were not displayed on the screen using a priming paradigm in the current L2 naming trials, the results indicate that the wrong member of the co-hyponym pairs was activated for the target word at some point of word recognition processing. This analysis suggests that semantically related and visually similar lexemes within the same lexical field tend to activate each other in an early stage of L2 Japanese word recognition. The finding has the pedagogical implication that early readers of kanji need more visual discriminatory power for the accurate word identification. Although the traditional kanji teaching has been pointed out this issue, the focus is often drawn on the common element more than on the different part of kanji, because the shared radicals give us useful semantic cues. Nevertheless, the current study revealed that early L2 readers notice such parts of the word easily, while other parts of kanji, which provide a vital discriminative cue, were not consciously noticed. From a practical point of view, a form-focused instruction (FFI) to draw learners' attention to a marked non-radical part of a kanji form, would encourage early readers to gain more powerful visual disambiguation skills. For example, in our sister-pair, the marked part of 妹 (*younger sister*) and 姉 (*elder sister*) should be the right part of non-radical constituent rather than the left-semantic radical. Similarly, the middle kanji morpheme should be focused in the vehicle co-hyponyms 自動車 (*car*) and 自転車 (*bicycle*) rather than the common initial and final kanji morphemes between the two lexemes.

Now, we turn to answer a question that was raised by the findings of previous chapters 3 and 4. Recall that we found a persistent L2 specific writing strategy, in which EJ writers often decompose a kanji compound into sub-lexical components. However, we could not so far determine any particular cause for such an uneconomical writing process. In the current study, we found evidence that low frequency Japanese words are likely to be read in a hit-or-miss manner by means of an unreliable sub-lexical matching strategy. The point is that the strategy itself is not at all a bad one, as L1 adult readers also employ the same strategy to read a never-encountered kanji compound. What is different between L1 and L2 reading is that native readers exploit it only as an emergency back-up strategy, while our EJ learners, with a lesser-developed lexical representation, heavily rely on this reading strategy, so that orthographic-to-phonological computation at a sub-lexical level becomes routine for them. However, another reading strategy (i.e. synonym readings) indicates that some lower proficiency readers are aware that sub-lexical matching strategy is not very effective.

In short, the two reading strategies found in the present study suggest that (1) EJ readers of kanji read less familiar kanji compounds by replacing them with a more frequently used familiar synonym; but (2) they have to employ the sub-lexical matching strategy for kanji compounds either with a skeletal underspecified representation or with no lexical entry at all, as they are unlikely to have a clue about the lexical meaning and its phonology. In addition, they have a sub-decomposing strategy at a morpheme level for unknown single kanji. The findings of this study provide evidence that the sub-lexical matching strategy is a by-product of low quality lexical representations or, in the worst case, they do not have lexical entries. The pedagogical implication of this position is that the learners need to be told at as early a stage of learning as possible that it is critical to acquire the whole word phonology of words with addressable orthographic representations if they want to become skilled reader/writers of Japanese. For beginning EJ learners of Japanese, the sub-lexical matching strategy works so well to decode unknown English words, so that it is necessary for them to know that their routine exploitations of the decomposing strategy to decode or encode Japanese words at sub-lexical level is often more harmful than helpful.

The spelling tasks

The next question regarding spelling tasks is whether native orthography (romaji) is strongly linked to phonological representations of EJ learners. In short, we did not find evidence for a positive L1 script effect of romaji regardless of proficiency level. For both proficiency groups, writing order had no effect on their writing performance. There could be several reasons for the lack of statistical significance in the current study; for example, I may not have had enough items or participants to produce differences. Nevertheless, the result suggests that the use of L1 script (romaji) did not encourage accuracy in transcribing L2 Japanese words. As Table 6.5 illustrates, the written production of participants from both groups frequently matched their pronunciation, regardless of whether the word is being spelled in romaji or hiragana. This suggests that our EJ learners start relying on L2 kana orthography at an early stage of their acquisition of Japanese literacy.

In Tamaoka and Menzel (1994)'s study, they found a positive L1 script effect for text reading (comprehension) and L2 pronunciation of Japanese with three alphabetic native speakers (German, French, and English). From Tamaoka and Menzel (1994)'s study, combined with the present study, it may be tempting to conclude that L1 script facilitates L2 'reading' in terms of text comprehension and pronunciation, but that there is no such effect on L2 'writing'. However, it is important to consider the difference in the number of participants in the two studies. While there was only one English participant in Tamaoka and Menzel (1994)'s study, the current study was conducted with twenty-nine English-speaking participants. Similarly, a recent training study of Hatasa (2002) also provided a contradictory result to Tamaoka and Menzel (1994)'s study. Hatasa (2002) reported that 8 weeks of time lag in the introduction of hiragana to American university students had no significant effect between the experimental group (i.e. lag group using romaji up until week 7) and the control group (i.e. no lag, using hiragana from the beginning). She concluded that the use of native orthography (romaji) did not affect early development of American students' reading ability, and that the lack of a significant difference indicated that hiragana and romaji had similar effects in L2 reading development Hatasa (2002, p.361). Thus, although it is difficult to compare the studies directly, because of the different methodologies, it is striking that the two different paradigms of Hatasa (2002)'s training study and the current approach taken by the present study produced a similar finding on the effect of the use of romaji in L2 reading and writing by English-speaking learners of Japanese. Mann (1986a)'s cross-linguistic study with Japanese and American children can

help us in understanding these outcomes. She noticed that learning the alphabet (romaji) in the Japanese education system takes place after 4 years of learning kana syllabary and kanji, and she investigated whether exposure to an alphabet brings about phoneme awareness in Japanese children. Mann reported that 4th graders were able to handle both syllable and phoneme manipulation tasks regardless of their alphabetic literacy. She interpreted that knowledge of an alphabet might not be the only determinant of phoneme awareness, but that any type of phonological orthography (in this case, kana syllabary) endorses this awareness. In this light, it is not surprising that hiragana and romaji had similar effects in reading (i.e. pronunciation tests) and writing (i.e. transcribing pronunciations) performance of English learners of Japanese, both in Hatasa's study and the present study.

Turning back to our original question now: is writing Japanese words in romaji easy for EJ writers of J-word processing? At an early stage of the current study, we speculated that writing the pronunciation of Japanese words with romaji using the QWERTY keyboard would be easier and faster for EJ learners, because of more efficient use of their L1 knowledge and the already acquired blind touch keyboard skill. This assumption now turns out to be a rather simplistic view. During the spelling task, I noticed that there were a number of participants, including novice writers, who automatically started to write the intended word in hiragana in the romaji order, or some participants started writing the word in romaji but unconsciously switched it to hiragana (e.g. 'kyoo しつ' for 'kyooshitu' *class room*) in the middle of their writing. It was an automatic processing that they were unconsciously writing and it was very hard to resist inhibition (i.e. do not write in hiragana).

Automatic processing requires a considerably long time practice, and it is rapid but less flexible once it becomes automatic (Shiffrin & Schneider, 1977; Norman & Shallice, 1986). This observation indicates that at least for some EJ learners, L2 orthography (hiragana) is more strongly linked to IL phonological representations than their native orthography romaji when they write by hand. This observation indicates that at least for some EJ learners, L2 orthography (hiragana) is more strongly linked to IL phonological representations than their native romaji orthography when they write by hand. This finding provides further evidence for Koda (1995)'s claim that early exposure to the L2 writing system is critical for the development of L2 literacy skills. As has been mentioned, our EJ novice learners learn hiragana from the beginning, so that by the year two, it would be possible to develop an automatic hiragana-sound connection like the one that we observed.

For L2 writers, the current study suggests that writing with J-word processor would impose on EJ writers a rather complicated mental operation. For one thing, L1 positive script transfer, if any, would be involved in a trade-off with their conscious use of hiragana convention rules. For example, a familiar word 'Tokyo' must be written as 'to-u-kyo-u' according to the hiragana convention or the 'to-kyo' (English) and 'to-o-kyo-o' (romaji convention) versions will end up with no kanji list. A further attentional load is that while typing a Japanese word with romaji under the hiragana convention, writers need to monitor alternative hiragana-kanji visual outcomes on the screen. Although they somehow manage to make themselves familiar with this perceptually demanding writing tool, an appropriate introduction and teaching materials, which cover key areas of perceptual, mechanical, and language difficulties, to L2 Japanese word-processing, would lift a considerable amount of burden from beginning L2 writers.

CHAPTER 7

L2 KANJI VERIFICATION PROCESS

7.1 L2 kanji verification process in discourse

7.1.1 Motivation and background

The purpose of this study is to investigate the L2 kanji verification process at the sentence level. The Japanese kanji reading literature generally favours the parallel-access view of kanji reading, in which both phonological and orthographical information are involved in the activation of the meaning of kanji words (see more detailed review in 2.3.3). This line of research provided evidence for two major effects in kanji reading (Sakuma et al., 1998). Firstly, in a semantic decision task, L1 readers of kanji failed to detect homophonic kanji significantly more often than non-homophonic kanji (i.e. homophony effect). Secondly, incorrect kanji that is visually similar to the target caused longer reaction time and higher error rates (i.e. visual similarity effect). Thirdly, the effect was strongest for the items where two factors were combined.

Interestingly, the first problem - the detection failure of homophonic kanji, is also widely recognized among L1 adult writers using Japanese word-processors. Although L1 writers, including professional writers and proof readers, are very careful not to overlook such homophonic kanji errors, we fairly often fail to notice them on-line, though the error is obvious once we spot it in the later editing stage or in the printed document. This phenomenon indicates that some kanji errors in word-processed documents are not simply the end-result of L1 writers' lack of kanji knowledge. Surprisingly, relevant L1 and L2 literature has never taken up this issue. However, recent findings in the field of visual perception could provide us with useful insights into how people perceive visual information. The field of visual

attention has especially challenged the traditional view that people's internal representation of the visual world is complete, detailed, and accurate (Simons & Levin, 1997; Rensink, O'Regan, & Clark, 1997; Simons & Chabris, 1999). For example, the 'Change Blindness' (CB) phenomenon has established that people are unable to notice very large changes to scenes even if the changes take place in full view. Rensink et al. (1997) provided evidence that CB phenomenon was not related to observers' eye movements. They used a novel method called the 'flicker' technique, in which a blank screen (flicker) was inserted briefly between the original and the modified pictures, and the participants were asked to look for any change in the modified picture. It was found that observers quickly and easily spotted changes under the no flicker condition, while under flicker condition people had enormous difficulty in locating the change, especially changes in a less focused part of the scene. Similarly, several related studies found that people often fail to notice the presence of unexpected objects (e.g. a faked gorilla in a basketball game) when their attentional focus was placed on a particular item, which is known as 'inattentional blindness' (IB) (Mack & Rock, 1998). The point to note is that the unnoticed items in both CB (changes) and IB (unattended objects) became obviously visible to the participants once they were mentioned. This line of research strongly suggests that limited attentional capacity is the key cause of the occurrence of the CB or IB phenomenon in visual field. An implication to the current study may be that resource-limited attentional effects are at work when visually and semantically distorted homophone kanji errors escape the notice of native writers, and even of professional proofreaders who always consciously search for such errors.

There is plenty of evidence in the semantic illusion literature to support the assumption that people often fail to notice a linguistic anomaly even if they are asked to detect distortions, and even if they hold the requisite real-word enencyclopaedic knowledge (Erickson & Mattson, 1981; Kamas, Reder, & Ayers, 1996; Oostendorp & Kok, 1990; Reder & Kusbit, 1991). For example, a question such as 'how many animals of each kind did Moses take on the ark?' induced the so-called 'Moses illusion' (Erickson & Mattson, 1981). That is, even though many people knew the fact that it was Noah who took the animals on the ark, people more often responded 'two' for the Moses-version than in an Adam-version, when they were asked the above question with alternative proper names X (e.g. Moses and Adam). In addition, as in CB and IB phenomenon in the visual field, semantic illusions are fairly robust, and people fall for illusions even when they are asked to detect semantic distortions. A similar line of research provides further evidence that anomaly detection is affected by global pragmatics. For example, Barton and Sanford (1993)

investigated how far the core meaning of the word 'survivors' is computed with different scenarios. They reported that 60% of participants detected the anomaly in the sentence - 'the authorities were trying to decide where to bury the survivors', which was embedded in a discourse about a crashed plane. Other versions such as 'When an aircraft crashes, where should the survivors be buried?' and 'When a bicycle accident occurs, where should the survivors be buried?' obtained 33% and 80% of detection rates respectively. Here, an anomaly in the less well fitting scenario of 'bicycle-crash' was detected much more often than in the more relevant 'survive-air-crash' context, in which an anomaly between the underlying presuppositions of 'survive' (need to be alive) and 'bury' (need to be dead)' were ignored at the local semantic level. A recent study of Sanford (2002) provided further evidence regarding attentional effects in lexical semantic processing, showing that linguistic focusing devices such as clefting, foregrounding, and focalization contribute to higher-level written text interpretation, in which readers only perceive centrally attended linguistic items while ignoring other less important parts of the text. The attended materials are processed in detail (i.e. "deep" processing), while less focused background information is processed in a "shallow" manner. In this way, Sanford (2002, p.203) claims, we control our attentional allocation, so that we can economize the enormity of inferential work, but because of this graded processing, lexical semantic processing in discourse is often incomplete or shallower than we normally assume. This view may offer the key answer to the question of what is the impact of visual similarity and contextual relatedness in our failure to notice kanji errors in sentences.

7.1.2   Research hypotheses

The study will examine the assumption that writers' kanji verification processes are affected by the degree of similarity of the error kanji to the target kanji, in terms of visual and contextual relatedness. I predict that the degree of visual similarity of the candidate kanji to the target kanji is first checked against internal representations, and if the overlap is high, then other things being equal, the verification process is less detailed than if the overlap is low. For visually similar kanji words, the discourse is checked. If the error-kanji is directly relevant to the sentence context, then it is more likely to be overlooked than an error kanji that does not fit well with the context situation. This yields the following research hypotheses:

1. Kanji non-words that are visually similar to the target kanji (V+) will be more frequently missed than with visual dissimilar kanji (V-) non-words.
2. There is no effect of contextual fit on visually dissimilar kanji (V-) non-words.
3. Visually similar kanji (V+) non-words in highly related contexts (C+) will be more frequently missed than visually similar kanji in less related contexts.

7.1.3 Method

Participants

This study included two groups (L1 and L2) of participants. For L2 participants, we first set the criterion that their reading skill should be at the advanced level of the Japanese Proficiency Grade (JPG) 2. Eighteen English-speaking learners of Japanese (Male=10, Female=8) at the University of Edinburgh participated in the study. They were given course credit for their participation. A further ten English-speaking volunteers (Male=4, Female=6) took part in the study. Twenty-six Japanese native speakers (male: N=6, female: N=20) aged between 19 and 40 years (M=25.1, SD=5.4), who had taken their formal education in Japan, took part for a small fee. They were attending an English course at the Institute for Applied Language Studies, the University of Edinburgh, and most of them had lived in the United Kingdom for less than 2 months (M=2.25 months) at the time of the experiment. Six additional native speakers of Japanese (male: N=1, female: N=5), who were post-graduate students of the University of Edinburgh, also volunteered as participants. Thus there were twenty-eight L2 participants and thirty-two L1 participants.

Materials

Twenty-four pairs of kanji (stimulus kanji words, and visually dissimilar homophonic kanji words of each stimuli) were used in this study, all of which were two-kanji compounds in the *on*-reading, with two to four morae in length. We first selected 150 words (two-kanji compounds) within the range of the Japanese proficiency Grade 2 level kanji words, from which twenty-four words were chosen as stimulus items, being selected for having a visually and semantically dissimilar homophone. We then created additional pairs of *non-word* two-kanji compounds by combining existing single kanji. The pair of pseudo-homophones shared identical phonological information to the real word stimuli, but the degree of their visual similarity to the target word was manipulated (i.e. visually similar, or visually dissimilar). The

definition of visual similarity of the kanji was set as follows. Visually dissimilar (V-) kanji were defined as (1) phonological information being identical, and (2) no overlap in their visual and semantic representation. Visually similar (V+) kanji was defined as (1) phonological information being identical, and (2) non-word kanji compounds partially overlapping in their visual representation in such way that the stimuli and the non-word share the first constituent kanji, and the visual information of the second constituent kanji partially overlapped, but there was no overlap in their semantic representation.

Test-1 (kanji error detection, see Appendix P) featured pairs of non-word two-kanji compounds in order to avoid the effect from kanji frequency that is associated with using existing words. A pair of short sentences with two contexts (biased condition: highly related context, neutral condition: less related context) was created for each kanji stimulus. The location (string position) of error kanji was identical in each context. In addition, the length of each sentence was also controlled (biased condition: M=10.7 words, neutral condition: M=10.8 words). Finally, four conditions of each stimulus were compiled: (1) visually similar, contextually related (V+C+), (2) visually similar, contextually not-related (V+C-), (3) visually dissimilar, contextually related (V-C+), and (4) visually dissimilar, contextually not-related (V-C-). Thus the study employed a 2x2 factorial design. An example material is given below in the four conditions of the stimulus 気候 ('kikoo' *climate*). The V+ non-word (気喉 'kikoo') and V- non-word (記広 'kikoo') are embedded in a short sentence '*We have recently talked about the climate that has particularly (C+) become warmer / (C-) being discussed*':

> V+ C+: 最近、特に温暖になった**気喉**について話しました。
> V+ C-：最近、特に話題になった**気喉**について話しました。
> V- C+: 最近、特に温暖になった**記広**について話しました。
> V- C-：最近、特に話題になった**記広**について話しました。

We completed a questionnaire test with thirty native teachers of Japanese prior to the experiment to see if the experimental materials (context) were sufficiently biased. In this test, each experimental sentence was given, but with the error-kanji replaced by the actual target kanji. The sentences were given in both the biased and neutral conditions, and the native speaker judges were asked use a 7-point scale to evaluate the predictability of the target kanji compound, which was printed in bold font, in the context of the sentence. It was expected that the sentences in the contextually related condition would be judged as more predictable

than those in the contextually unrelated conditions, which would show that the contextual manipulation was sufficiently strong. To evaluate the Japanese judgers' scores on the contextual relatedness judgment, a paired-sample t-test was conducted on the mean scores for each participant, collapsed over items (t1), and on the mean scores for each item, collapsed over participants (t2). There was a significant difference between the judgments for the biased context (M=5.28), and neutral context (M=2.51) (t1(13)=13.7, $p < .0001$; t2(23)=11.07, $p < .0001$). Thus, we concluded that materials were sufficiently biased.

Finally, a further forty-eight sentences (12 with errors, 36 without errors) were created as fillers. The location (string position) of errors among twelve fillers was distributed between the sentence initial kanji (N=5), the second character in the sentence (N=2), and the sentence final kanji (N=5). The level of every kanji in the 72 sentences was checked.[1] Overall, the Kanji in experimental sentences included 90.3% grade-2 or lower level kanji, and 96.6% of kanji in filler sentences were grade-2 or lower levels. The 72 sentences were randomized. Four experimental files were created, each of which contained only one of the four conditions for any given experimental item. Overall, each file contained an equal number of items in each condition.

Test-2 (see Appendix Q) was designed to assess actual kanji knowledge for each participant. The sentence was given, including a blank where the critical kanji compound had appeared. Below the sentence, four versions of the target stimulus (correct kanji compound, V+ non-word, V- non-word, and homophonic kanji compound) were given in a random order, and the participants were asked to circle the right kanji in the list. They were also asked to circle a number underneath each item, representing how certain they were about their answer. The numbers ranged from 1 to 5, where 1 represents completely unsure, and 5 represents completely certain, and the numbers in the middle represented degrees of certainty between these extremes. For example, four versions of 'kikoo': a) correct kanji compound, b) V+ non-word, c) V- non-word, and d) homophonic kanji word were given for the underlined part of the sentence '*We have recently talked about the climate that has particularly become warmer*' below:

最近、特に温暖になった（　　　　）について話しました。
a) 気候　　b) 気喉　　　c) 記広　　d) 帰港
Completely unsure　　1　　2　　3　　4　　5　　*Completely certain*

---

[1]An on-line pedagogical tool 'Reading Tutor' at http://language.tiu.ac.jp/tools.html provided this information.

Test-2 is important because it allows us to discount those detection failures that are simply the result of a lack of kanji knowledge. In this study, what we are interested in is cases where participants fail to notice the kanji error in Test 1 *despite* having the requisite knowledge.

Task and procedure

Most of the participants were tested as a group in a quiet room. Participants were randomly assigned to one of the four experimental files. They were asked to read all 72 sentences within a given time, which was set differently for L1 (maximum 10 minutes) and L2 (maximum 20 minutes) readers. Immediately before the experiment we instructed participants to read the printed sentences as quickly as possible (speed-reading), and to cross out any kanji that they thought was incorrect. They were also instructed not to go back to previous sentences once they had done the task. They were not allowed to read the sentences in advance, to use a dictionary, or to confer with others about the task. After a break, Test-2 was conducted on the same day. We began the second test with the instruction to circle the right kanji in the list as well as circling a number underneath each item, representing how certain they were about their choice. There was no time limit for the second test.

7.2 Results and discussion

7.2.1 Overall analyses

The present study examined how writers' kanji verification processes were affected by the visual similarity of error kanji and contextual relatedness. The first question was whether V+ (visually similar kanji) kanji errors were more frequently missed than with V- (visual dissimilar) kanji errors. The second question was whether V+ kanji errors in highly related contexts (C+) were more frequently missed than V+ kanji errors in less related contexts. Each question will be answered separately.

Figure 7.1 shows the overall performance of kanji error detection in Test-1, by both L1 and L2 groups. While L2 kanji readers detected only 31% of the kanji errors, L1 readers detected 89.1% of the kanji errors.

The most likely explanation for this large difference in error detection may be that our experimental materials were chosen from the range of the Japanese proficiency

Figure 7.1: Detection rate by group
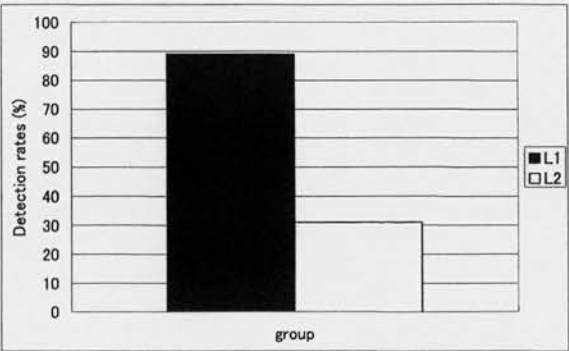
Grade 2 level kanji words, a standard of which was especially set for the L2 learners of Japanese. Therefore, it was obviously an easy task for L1 readers.[2] However, a further analysis revealed that L1 readers did not detect every kanji error consistently. Figure 7.2 shows detection rate of each condition by L1 group, and Figure 7.3 shows the results for the L2 group.

Figure 7.2: Overall percentage detection rates (L1)
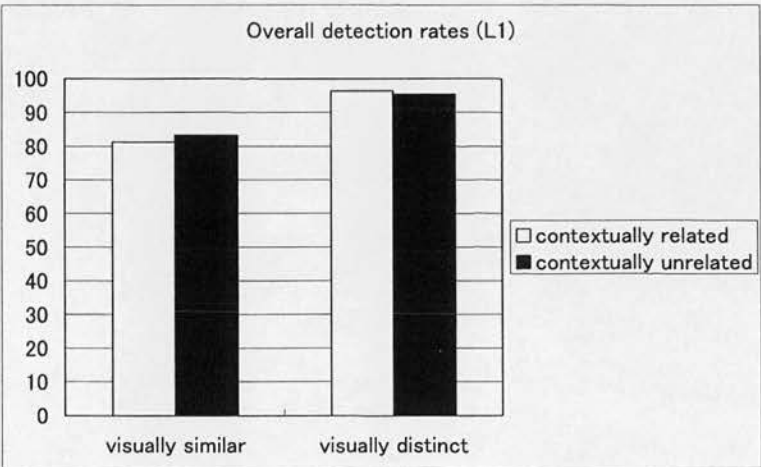
The results of the L1 kanji error detection numerically indicate an effect of visual similarity that were similar to those of Horodeck (1987), Matsunaga (1995)

---

[2]Most of L1 participants were undergraduate students or they had at least undergraduate level education.

Figure 7.3: Overall percentage detection rates (L2)
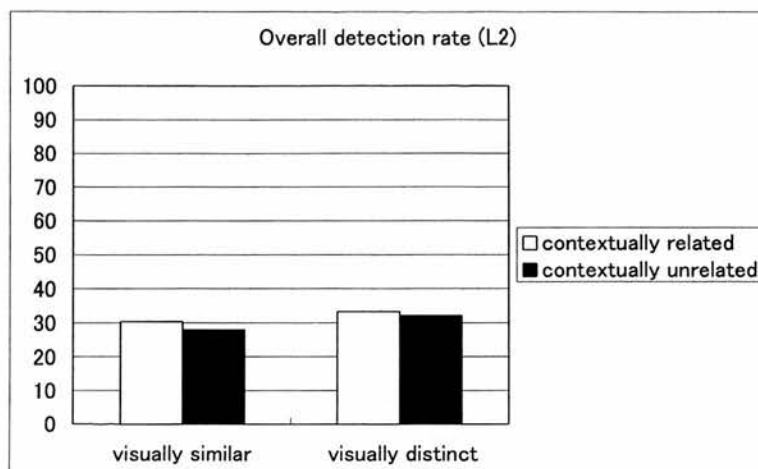
(kanji error detection within context), and other previous studies of Sakuma et al. (1998), Wydell et al. (1993) (kanji error detection in isolation). L1 readers in the present study missed more visually similar kanji non-words (M=82.3%, SD=17) than visually dissimilar kanji non-words (M=95.8%, SD=8.4) to the target. Two-way analyses of variance confirmed the visual similarity effect among L1 readers ($F1(1, 31)=30.64$, Mse=0.019, P < .05; $F2(1, 23)=14.78$, Mse=0.0298, P < .05). The effects of contextual similarity and interaction between visual and contextual similarities were not significant, however (all F's < 1). It is clear that L1 readers recognized visually (semantically) anomalous kanji compounds easily upon reading them. Thus, we can reasonably conclude that L1 readers checked the degree of visual similarity of the candidate kanji to the target kanji, and the visually similar kanji escaped notice from the L1 readers' verification process much more than visually distinct kanji.

On the contrary, although L2 readers recognized numerically more visually distinct kanji (M=32.7%, SD=30) than visually similar kanji (M=29.2%, SD=24.5), the visual similarity factor had very little effect on EJ kanji readers (both F's < 1). The main effect of contextual relatedness and two-way interaction were likewise non-significant (p's > .2). In other words, L2 readers often did not notice a local anomaly, even when it was cued by a completely different shape of kanji to the target kanji, embedded into an otherwise meaningful sentence. In short, the L2 data shows very little sign of the widely acknowledged visual similarity effect.

7.2.2   Unnoticed Kanji errors by L1 and L2 readers

In accordance with findings in the visual attention and semantic illusion literature, we hypothesized that resource-limited attentional effects were at work when visually and semantically distorted homophone kanji errors escape from eyes of L1 writers who always consciously search for such errors. The key point here is that people often fail to notice visual or linguistic anomalies even if they hold the right knowledge, and people fall for illusions even when they were asked to detect anomalies or changes.

The previous section established an effect of visual similarity in the proportion of kanji detection errors. However, it is important to establish that these effects are due to genuine failures to notice the errors, rather than, for example, to missing lexical knowledge. In this section, we will therefore consider only those trials where the participant demonstrated correct kanji knowledge on Test-2. We defined a 'unnoticed kanji error' as a trial in which the reader failed to detect a wrong kanji in the experiment even though he or she subsequently demonstrated the correct lexical knowledge of the given kanji (i.e. identified the correct kanji in test-2). Then, each participant's *unnoticed kanji error rate* was defined by taking all trials on which the participant correctly identified the appropriate kanji in test-2, and calculating the proportion of these trials in which the participant failed to identify the corresponding non-word in test-1. The phenomenon of unnoticed errors can be argued to be similar to that of semantic illusions cited above: in both cases, there is a failure to detect an anomaly despite the participant's knowledge. Another way of looking at this is in terms of test-retest reliability; the 'unnoticed kanji errors' correspond to cases where the participant fails on an initial test item, but succeeds on the same item in a re-test. However, both the aims and design of the current study are different from those of a test re-test design, where the researcher is usually interested in testing the reliability of a test over time, by correlating results between two applications of the test (see Hatch & Lazaraton, 1991, p. 531). In contrast, in the current study, the main point of interest lies in the actual *differences* between the two tests. Furthermore, the tasks in the two tests differ, so that Test-1 (detection of kanji errors embedded in sentences) is designed to elicit errors, while Test-2 (multiple choice) is designed to detect the participants' correct knowledge. Table 7.1 gives a summary of the 'unnoticed kanji error' phenomenon in both groups, and Table 7.2 gives a break-down of the distribution of unnoticed kanji errors by condition:

Table 7.1: Summary of unnoticed kanji errors

	L1 test-2 correct (n=733/768)	L2 test-2 correct (n=450/672)
Test-1 undetected kanji error	82 out of 733 (11.2%)	282 out of 450 (62.7%)
Test-1 non-word identified	651 out of 733 (88.8%)	168 out of 450 (37.3%)

Table 7.2: Distribution of rates of unnoticed kanji errors between conditions for the two participant groups

	L1	L2
visually similar/context related	34 out of 182 (18.7%)	79 out of 122 (64.8%)
visually similar/context unreltaed	32 out of 183 (17.5%)	79 out of 116 (68.1%)
visually distinct/contcxt rclated	7 out of 186 (3.8%)	58 out of 104 (55.8%)
visually distinct/context unrelated	9 out of 182 (4.9%)	66 out of 108 (61.1%)
Totals	82 out of 733	282 out of 450

The average per-subject means for each condition are given in Figures 7.4 and 7.5[3]

According to their meta-evaluation about the degree of certainty on their selected kanji, L2 readers were certain (in 3-5 certainty range of the scale) of 67% (n=189) of their answers when they identified the right kanji, while the remaining 33% (n=93) of the correctly selected kanji fell in the 1-2 completely uncertain range, that indicating either that L2 readers underestimated their correct lexical knowledge or that there were some chance hits. However, all 28 of the L2 participants scored well above the chance level of 25% for the 4-way multiple choice task of test-2 (Mean=67%, SE=3.9, binomial sign test, $p < .001$). Thus, we can safely conclude that very high L2 error detection failure rate was not mainly due to their lack of kanji knowledge. However, the L2 data summarised in Figure 7.5 suggests that neither visual similarity nor contextual relatedness clearly affected on their rates of unnoticed errors.

Two-way analyses of variance were performed on the unnoticed error rates for L1 and L2 readers. The first ANOVA analyzing unnoticed error rates in L1 data confirmed a visual similarity effect among L1 readers with a significant result (F1 (1, 31)=27.95, Mse=212.39, p< .05; F2(1, 23)=13.56, Mse=311.05, p< .05). The effects of contextual similarity and interaction between visual and contextual similarities were not significant (all F's < 1). Although there was a marginal main effect of visual similarity in the items analysis of the L2 data (F2(1,23) = 3.64, MSe

---

[3]Note that because these are averages of subject means, the percentages may differ slightly from those in Table 7.2, which shows data collapsed over subjects.

Figure 7.4: Unnoticed Kanji errors (L1): Averages of per-subject means for each condition
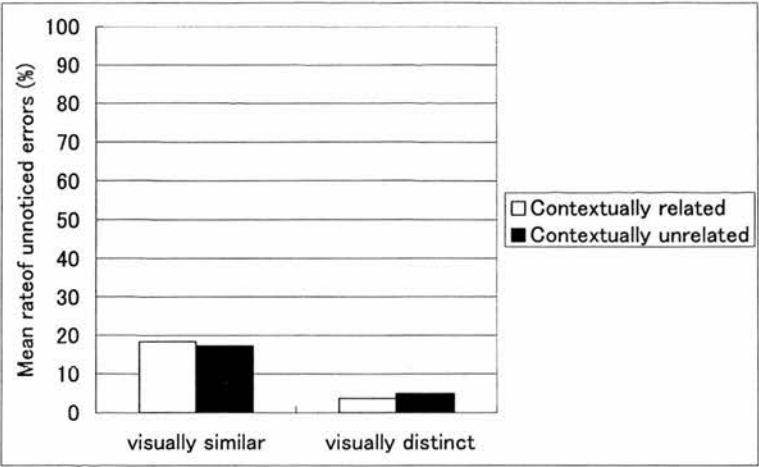
Figure 7.5: Unnoticed Kanji errors (L2): Averages of per-subject means for each condition
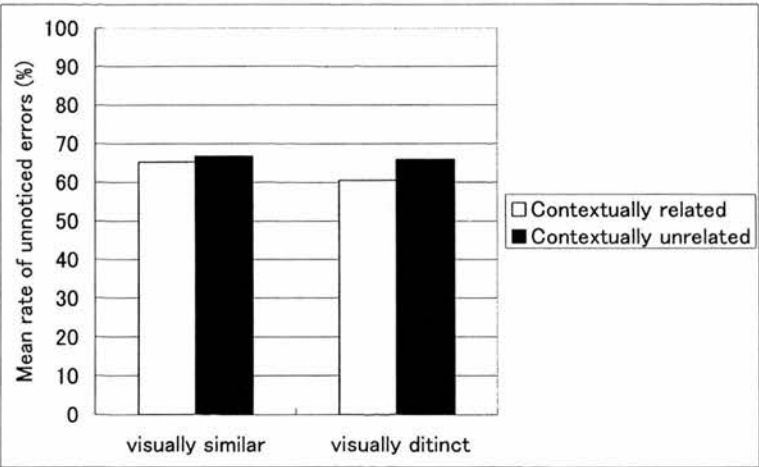
$= 1734$, p$<$ .07) no other effects approached significance (all p's $>$ .2).[4] Thus, it was confirmed that visual similarity had effect on L1 readers only, and contextual relatedness has no effect on unnoticed kanji error rates for both L1 and L2 readers.

One explanation for the lack of statistical significance for contextual factor may be that our experimental sentences were not biased enough for the two types of

---

[4]Because of missing data, one design cell had to be replaced in the items analysis for the L2 data. The replacement procedure was to replace the cell with [ItemMean + ConditionMean - GrandMean] (Winer, 1971).

sentence contexts. However, this is unlikely to be the case, since we statistically confirmed that there was an experimentally detectable difference before running the experiments. Another possibility could be that the L1 readers concentrated on finding visually anomalous kanji (i.e. error monitoring) and did not read the kanji for meaning (i.e. true reading). The same question was raised by Matsunaga (1995) for Horodeck (1987)'s L1 kanji study, but Matsunaga (see more detailed discussion in section 2.3.3) concluded that L1 adult readers genuinely read target short sentences for meaning. Although the kanji used in Horodeck's L1 study and the present study were not identical, the reading condition (i.e. reading short sentences as quickly as possible) was very similar, and most of us would agree that we cannot help processing word meanings while reading a meaningful sentence. The most likely explanation would be that a strong combined effect of homophony and visual similarity made L1 verification process less rigid for visually similar (v+) items even in sentences with less relevant context once they had gone through the first check against visual representation.

L2 kanji readers' visual recognition skills under a time-pressure were not good enough to detect either visually similar kanji errors, or even visually less-overlapped kanji errors, that were not predictable from the surrounding context. The results for L2 readers in the present study support recent L2 kanji studies of Mori (2002, 2003) that not all L2 readers benefit from the availability of multiple sources of contextual and morphological information, as almost half of the L2 readers (American college students) in that study over-relied on one source (i.e. kanji or contextual clues). Mori (2003, p.411) reported that her American learners of Japanese made more semantically related guesses from kanji clues than from context, and they used more contextual clues for guessing syntactic information. Purely contextual aid for guessing the meaning of kanji compound words did not show any statistical significant results. Mori (2003) concluded that it was not easy for L2 American readers to guess the lexical meaning of unknown kanji words from even combined information of kanji and context.

There was, however, evidence in the present study that L2 readers were affected by visual similarity in test-2. We examined the L2 reader's wrong responses for the multiple choice test, and it was found that the visually similar non-word accounted for around twice as many of the errors as the other two error types. Table 7.3 shows this, adding a comparison with the L1 error data:

Table 7.3: Total number of errors for L2 and L1 multiple-choice task (test-2) by error type

	L2 error types	L1 error types
Test-2 visually similar non-word	110 (49.5%)	11 (31.4%)
Test-2 visually distinct non-word	53 (23.9%)	2 (5.7%)
Test-2 homophone (visually distinct)	59 (26.6%)	22 (62.9%)

A Friedman Chi-square test (non-parametric within-subjects analysis of variance) confirmed that the L2 pattern was significant across the participants ($\chi^2(2) = 12.02$, $N = 28$, $p < .005$). The L1 pattern shows a similar trend for visual similarity, but with a larger number of homophone errors (which was probably because the homophones were real words). However, we did not statistically analyze the L1 error types, because of the sparseness of the error data distributed across the thirty-two L1 participants. In any case, the pattern of errors suggests that the L2 participants were sensitive to the visual similarity of kanji, at least when they were engaged in the less demanding multiple-choice task, which was conducted without time pressure.

Chapters 3 and 4, which investigated written errors appearing in a meaningful sentence, found evidence for homophone kanji errors. In addition, we found an L2 specific error pattern in non-homophonic kanji errors, which included both real words and non-words. Reflection on the previous literature and the findings from the Chapter 4 as well as the present study make it clear why EJ writers faced a lexical judgment problem of whether a given kanji compound was a legitimate kanji word (existing word) or not (non-word). First of all, we identified an L2-specific sub-word matching strategy, and interpreted it as a by-product of low quality lexical representations. In the previous chapter, we provided evidence that it was very hard for L2 readers to guess the whole word pronunciation of a kanji compound by matching the sub-word phonology of each constituent kanji. Likewise, Mori (2003)'s study provided evidence for a similar L2 sub-word semantic assembly problem for kanji compounds. In short, we cannot expect L2 kanji readers to perform a native-like intuitive judgment about the lexicality of a given kanji compound. Thus it is possible that EJ readers in the present study regarded visually dissimilar (V-) kanji items as unknown real words rather than non-word homophone kanji. Secondly, the high unnoticed kanji error rate among EJ readers in the present study suggests the shallow nature of the L2 kanji verification process. For example, in L2 reading, many visually dissimilar kanji were not ruled out upon reading them, as if they had slipped through a loose net. Thus it is perhaps right to say that global pragmatics

has much less influence in L2 anomaly detection, and L2 lexical-visual/semantic processing was very shallow during the kanji verification process. If we can extend this finding to L2 kanji reading and writing in normal situations, including in J-word processing writing environments, it explains why EJ writers often unnoticed a local semantic anomaly of kanji that is deeply embedded into background context, rather than it visually sticking out from the surrounding context. To summarise, the findings in the present study further support Mori's claim (2002) that teachers should encourage more use of multiple information sources in L2 vocabulary learning, such as the use of contextual information and kanji information, so that L2 learners of Japanese will be equipped with a better kanji decoding strategy.

CHAPTER 8

CONCLUSIONS

8.1 Overall Summary of findings

8.1.1 Error Analysis

Chapter 3 was concerned with writing (spelling) problems encountered by English-speaking (EJ) learners of Japanese when writing with Japanese word processors. Error analysis (N=340) revealed the following two major error categories:

Firstly, as an L2 specific phenomenon, we observed the appearance of orthographically illegal hiragana strings with a partial deviation from the correct phonological input. This type of hiragana error was found to constitute a major category in our EJ corpus. This finding provides evidence that the L2 writers' kana literacy level plays an important role when they write with J-word processors. I further speculated that the structure of the kana syllabary was in some way reflected in L2 word-processed spelling errors in Japanese. Consequently, Chapter 5 (Moraic Awareness study) was designed to tap the internal structure of the Japanese syllable among EJ learners of Japanese at different proficiency levels, which in return provided a perceptual baseline data for further chapters.

Secondly, we observed that the well-known L1 phenomenon of homophone kanji errors was also found in EJ kanji errors. However, I found that there were more non-homophonic errors than homophonic kanji errors in our EJ corpus. Various lines of evidence strongly suggest that EJ specific kanji error patterns in non-homophonic kanji errors involved multiple errors mainly based on phonological and visual factors. These kanji errors suggested that many L2 writers had error detection problems, so that such heavily distorted non-homophonic kanji errors escaped while learners

152

were performing on-line or/and off-line kanji verification. A question was raised as to whether these errors occurred as a result of a perceptual problem (e.g. detection failed because of weak visual recognition skills) or linguistic problems (e.g. detection failed due to the lack of a lexical entry). I thus addressed this question further in Chapters 4 (protocol analysis) and 7 (experimental study).

8.1.2 Factors in L2 keyboard spelling errors

Chapter 4 presented a preliminary investigation of the L2 writing process of J-word processor use, through observations and analyses of verbal protocol data. The purpose of this study was to determine how L2 writers write with J-word processors, and why and when word-processed spelling errors occur. Hence, the real-time screen movements were also recorded to observe physical writing behaviour such as the number of repair units. In order to narrow down the factors relating to mechanical errors and writers' keyboard skills, a first study (hiragana copy task) was carried out. The second study (hiragana production task) investigated how L2 writers utilize their lexical knowledge as an input to J-word processing. The third study looked into the way L2 writers conducted their kanji error monitoring tasks.

Keyboard skills The first study presented three pilot findings, which shed light on the L2 specific orthographically illegal hiragana strings that were identified in Chapter 3. First, I found L2 hiragana visual recognition problems where L2 novice writers' visual recognition processes were affected by surrounding context. Secondly, I found evidence that writing an orthographically ambiguous word caused a writing problem. That was that L2 writers tended to rewrite more when they wrote an orthographically ambiguous word, including hiragana options that share the same phonetic value. Thirdly, there was a typing problem regarding geminate consonants, for which the writer always needed to hit the same key twice.

Retrieval memory and knowledge of orthographic conventions The production task revealed three L2 specific phenomena:

1. Errors concerning addition or omission of geminate consonants seemed to be triggered by a combination of grammatical formation problems and slips.

2. A sub-lexical writing strategy (i.e. decomposed a kanji compound word on purpose) was observed.

3. Long vowels /uu/ and /oo/ in word-final position seem to have been more problematic than other vowels.

Thus, this part of the hiragana production task generated a research question about whether deletion errors of the long vowel /oo/ in L2 written production would be a result of the underlying phonological representation or inadequate orthography acquisition. The design of experimental materials for Chapter 6 was based on the third finding in this section.

Kanji error monitoring skill   The third study suggested strategic differences between L1 and L2 participants in their problem-solving tasks, in which participants were asked to detect kanji errors, and these were then traced back to the target kanji compound word. The following results were obtained:

1. L2 subjects showed a tendency to make false-alarms by wrongly classifying a given kanji compound as a non-word (partial kanji recognition problem).

2. L2 subjects often employed a trial-and-error strategy, whereas L1 subjects' judgments were almost automatic.

3. while L1 subjects used both bottom-up and top-down strategies freely in the search for the right kanji, L2 subjects were less able to deal with the task with a holistic approach, and they handled the task locally at phonological, semantic or visual level of information for a given kanji.

Observations and protocol analysis in this section suggested a possible involvement of multiple factors besides learners' kanji knowledge or lack of attention when L2 kanji readers failed to notice kanji errors. Through the review of the literature, I predicted that such factors could be the effects of phonological and visual similarity and contextual relatedness. Thus, the findings in this section generated the research hypothesis for Chapter 7 (see below).

8.1.3   Moraic Awareness by L2 learners

Chapter 5 investigated the level of perceptual sensitivity to the internal structure of the Japanese syllable among EJ learners of Japanese at different proficiency levels.

Three groups of learners (monolingual, novice and semi-bilingual) were compared on the 'phonological counting task', which required listeners to tap the number of sound units. I sought to determine the extent of the use of L1-specific listening strategies, and in particular, whether EJ learners used the syllable or mora as the basic perceptual unit of segmentation. The following two findings were made:

1. The devoicing vowel environment led to a decrease in the perceived number of rhythmic units for the novice and monolingual groups, but the semi-bilingual group perceived morae identically, whether they had a voiced or a voiceless vowel.

2. For syllables including more than one mora, the different learner groups differed significantly in their perceived number of units. L1-specific listening strategies were strong among English speakers with none or very low exposure to Japanese language, while more experienced fluent learners' moraic awareness was found to be near native level.

Thus, I conclude that perceptual constraints are imposed through an L1-specific listening strategy, and the lack of psychological reality of the mora constituted a major challenge for the perception of special syllables for less proficient English learners of Japanese. The major finding of this study was that the semi-bilingual group, whose authentic language exposure was longest, was more influenced by the number of morae than the number of syllables in Japanese words. I strongly suspect that a ceiling level of kana literacy, which should be very similar to L1 adult Japanese among participants in this group, was highly responsible for this outcome, as the level of literacy is generally considered to play a key role in phonological awareness tasks such as the one we conducted.

8.1.4   Reading and spelling of Japanese words

In Chapter 6, I investigated EJ learners' interlanguage lexical representations in reading and spelling. This chapter consisted of two reading tasks (the pronunciation task and the naming task), and a spelling task. I posed the general question of whether a writer's linguistic problems are reflected in their written products, and we were especially concerned with the nature of the relationship between phonological and orthographical representations in terms of the loss of a mora in the long vowel /o:/.

The reading tasks

The reading tasks have produced three key findings. To begin with, I found evidence that spelling and reading share the same lexical representation. This finding supports previous L1 studies that hold a single lexicon view of reading and spelling. However, our L2 data showed clear individual differences. For example, some skilled readers still had a graphic representation deficit despite having perfectly acquired the relevant phonological knowledge, while some of the low proficiency readers developed a higher rate of perfectly agreeing phonology-orthography lexical representations, with a very limited number of acquired items. Thus, I have concluded that the phonology-orthography agreement rate did not increase according to the proficiency continuum. In addition, two types of lexical acquisition orders were identified. The first development pattern *phonology-then-orthography* acquisition coincides with a typical L1 development pattern. I have acknowledged the second development pattern *synchronized phonology-orthography* acquisition as the result of a typical L2 lexical learning situation (i.e. spoken and written forms of a given word are often taught simultaneously).

Next, I found several lines of evidence for the frequency effect, which is well documented in L1 research, in an L2 naming task. Since the SLA literature has just begun exploring the frequency effect on L2 reading, this finding suggest a potential interest in the use of Psycholinguistic-oriented methodologies for SLA research, such as measuring Reaction times of L2 readers.

Finally, I identified two L2 kanji reading strategies:

1. EJ readers of kanji read less familiar kanji compounds by replacing them with a more frequently used familiar synonyms; but
2. they have to employ the sub-lexical matching strategy for kanji compounds. In addition, they have a sub-decomposing strategy at a morpheme level for unknown single kanji.

Regarding the first kanji reading strategy, this part of the study also provided several lines of evidence that both phonological and semantic information played a role in the process of L2 word recognition. This was illustrated by an inappropriate identification of the wrong member of the co-hyponym pairs where phonological/visual and semantic relatedness was evident. The second reading strategy demonstrated that L2 readers were unlikely to have a clue about the lexical meaning and its phonology when they used this sub-lexical reading strategy, thus I interpret the

results as showing that the sub-lexical matching strategy is a by-product of low quality lexical representations or, in the worst case, they do not have lexical entries.

The spelling tasks

This task aimed to investigate whether native orthography (romaji) was strongly linked to phonological representations of EJ learners. I hypothesized that writing a word in romaji before hiragana should reduce errors in hiragana, if the difficulty is orthographic, but that if the difficulty is phonological, then writing in romaji first should make no difference to the hiragana errors. Our results showed that the written production of participants from both novice and skilled groups frequently matched their pronunciation, regardless of whether the word is being spelled in romaji or hiragana, and the writing order had no significant effect on their writing performance for both proficiency groups. In short, we did not find evidence for a positive L1 script effect of romaji regardless of proficiency level. Moreover, I reported that during the spelling task, some participants automatically started to write the intended word in hiragana when they should have been writing in romaji, or some of them unconsciously switched to hiragana in the middle of writing in romaji, while I did not notice the reverse situation. This observation suggests that L2 orthography (hiragana) would be more strongly linked to IL phonological representations than their native romaji orthography, at least for some EJ learners. As far as this study is concerned, I have concluded that the use of the L1 script (romaji) did not encourage accuracy in transcribing L2 Japanese words.

This part of study investigated another issue of spelling (hiragana) ambiguity in Japanese, since previous chapters 3 and 4 identified some EJ spelling errors that pointed to an influence of Japanese orthography regarding confusion between two-to-one hiragana-sound correspondences. We found counter evidence against the general belief among literacy educators of Japanese that L2 spelling difficulties concerning a choice between hiragana letters う-お option for the second element of long vowel /o:/ was an early learners' problem. This finding again points to the need of reevaluation of the role of kana literacy in L2 orthography acquisition.

8.1.5 L2 kanji verification process in discourse

Chapter 7 examined from the perspectives of Psycholinguistics how L2 writers' kanji verification processes were affected by the visual similarity of error kanji and contextual relatedness. The learner factor (i.e. individual learner difference) is one of key issues in an SLA research. Thus, this study measured an L2 'kanji illusion' where the reader failed to detect a wrong kanji in the experiment even though he or she subsequently demonstrated the correct lexical knowledge of the given kanji.

I found clear-cut evidence for a visual similarity effect on L1 readers who recognized visually (semantically) anomalous kanji compounds more easily than they recognised anomalous kanji compounds that were visually similar to the target. On the other hand, L2 kanji readers' visual recognition skills under a time-pressure were not good enough to detect either visually similar and dissimilar kanji errors. However, in the less time demanding multiple-choice test, I found evidence that L2 readers were obviously affected by visual similarity. According to their meta-evaluation about the degree of certainty on their selected kanji, L2 readers were certain of nearly 70% of their answers when they identified the right kanji, and their choice of correct kanji was found to be beyond chance. Thus, I conclude that the very high L2 error detection failure rate was not mainly due to their lack of kanji knowledge. This finding suggests that holding correct semantic and visual lexical representations of a kanji compound does not guarantee a successful L2 kanji verification process. Indeed, the high rate of L2 kanji illusion suggests the shallow nature of the L2 kanji verification process, where global pragmatics has much less influence on L2 visual recognition.

8.2 Implications for L2 acquisition of Japanese orthography

8.2.1 On kana syllabaries and romaji teaching

Our result in chapter 7 suggests that there is no difference in the accuracy of L2 spelling, regardless of whether the words are written in romaji or hiragana. Similar results were also found in Hatasa (2002)'s training study, which found that the early introduction of hiragana had little impact on beginning L2 reading development. As Mann (1986a)'s L1 cross-linguistic study suggests, phonological orthography such as romaji and hiragana have had similar impacts on L2 reading and writing performance. However, our observation on L2 handwriting has demonstrated a possibly stronger connection between IL phonology and L2 hiragana over L1 romaji.

Naturally, I shall tentatively suggest that the early exposure to one of the L2 writing systems, hiragana, would facilitate the development of L2 writing skills. Here, I can only make a weak claim, since opinions are divided among L2 researchers and practitioners, and I have not tested this particular hypothesis in this thesis. However, the findings of this thesis imply that the teachers of Japanese should not consider the mastery of reading and writing of hiragana itself as the completion of hiragana literacy acquisition. It is not too difficult to encourage more use of hiragana during reading and writing classes. For example, more effective utilization of the kanji pronunciation indicator *furigana*, which is written in hiragana, would help L2 acquisition of the dictionary presentation of the whole-word sound of a kanji compound.

As for the electronic classroom in Japanese, the key issue seems to be the romaji input method. As we can see from Appendences B and C of this thesis, one of the problems is lack of unity in romaji usage. Thus, the first task for the instructor should be checking the romaji input system of a given writing tool, and utilizing a mora-romaji mapping list similar to the kana table shown in Appendix A. This would prevent L2 novice writers from making the kind of simple typing mistakes that were illustrated in section 4.1.

One area of Japanese orthography that I have largely avoided in this thesis is the katakana syllabary. This does not mean at all that katakana has a lesser role in the Japanese language, nor in the study of Japanese SLA. Far from it, there is much intriguing L2 research in the current literature on the relation between IL phonology and loan words, which are conventionally written using katakana letters, and some studies also extend their study to katakana orthography conventions (Bunkacho, 1971; Endo, 1989; Inagaki, 1991; Kobayashi, Quakenbush, & Fukada, 1991; Tamamura, 1991; Toda, 1999; Quackenbush & Oso, 1990; Quackenbush & Toda, 1992). The main reason for my decision not to include such an interesting research topic is that the orthographic conventions of katakana transcription of loan words is not rigid nor well-defined, due to the gap between Japanese phonology and the phonology of the original foreign words. For example, the National Language Council has allowed the use of the katakana sequence ヴァ /va/ to transcribe foreign words including the initial letter '*v*', such as *violin*, but because the old version バ /ba/ is still publicly in use, there are two permissible katakana transcriptions of *violin* (i.e. バイオリン and ヴァイオリン). The core problem is that native Japanese writers do not use the ヴァ /va/ and バ /ba/ versions in a consistent manner. For example, some L1 Japanese may freely use both the ヴァ /va/-version and the バ /ba/-version

when transcribing *Van Gogh*, while they seem somewhat consistent in using the バ /ba/-version for the katakana transcription of *van* (an estate car). Because I noted more than 30 such inconsistent katakana sequences, and 80.8% of the loan words in Japanese are of English origin (see, Shibatani (1990, p.148)), I predicted that too many confounding variables would be found with the English-speaking learners of Japanese who participated in this thesis. For this reason, this area of investigation has been left open for the future study.

8.2.2 Answers to three frequently asked questions on the use of keyboard-based writing tools

How can L2 learners write with the keyboard without making any mistakes?

The answer to this question summarises the implications of the thesis for the place of digital literacy tools in pedagogy, as well as computer-writing instruction in Japanese, L2 vocabulary development and the corresponding learning strategy.

Before giving my answer to this question, I need to touch upon two important issues regarding digital writing. First of all, I cannot give my answer without knowing the exact specifications of a given writing technology (i.e. what the given machine can offer and cannot offer to its writers) because as Hill, Wallace, and Haas (1991) accurately point out, "generalizing from results obtained with one interface to make assertions about computers in general may be unwise". Secondly, the rapid advancement of technology may one-day make it possible for L2 writers to produce a polished Japanese text without experiencing the spelling problems illustrated by this thesis. If an L2 learner of Japanese wishes to use a state-of-the-art digital literacy tool in order to write an academically sound essay in Japanese with the minimal possible effort, by concentrating on the content of the essay rather than writing process itself, I would be happy to accept this position.

But if another L2 learner believes that learning from own mistakes is useful and important as part of the language learning process, here is my suggestion. As for instructional matters, the problem would be to decide whether to change the learners themselves or to try to improve instruction. In the SLA literature, there is fairly general agreement that learner's learning strategies help their language learning to some extent, but research in this field has not yet provided empirical evidence to support the causal relationship between the two (Koyanagi, 2001). In this thesis, I demonstrated that linguistic and perceptual constraints were associated with some learner strategies, and the idea of changing the learners themselves would

imply an effort to remove such constraints. This is all very well and good, but the main problem of this approach may be that improving a learner's proficiency level is an extremely time-consuming task. Thus, it would be more plausible to consider improving writing instruction, with the aim to minimize the effect of known underlying factors of perceptual, linguistic, mechanical, and cognitive difficulties, while at the same time monitoring individual learners' progress in the long term.

As for the learners' part of the responsibility, the results of chapters 5 and 7 suggest that they need to make sure to achieve the ceiling level of hiragana literacy at as early a stage of learning as possible. In addition, once they have learned the spoken forms, they must also be sure to understand how these units of phonology are written in hiragana in a dictionary representation. In other words, they need conscious learning of lexical items in terms of sound unit-grapheme (mora-hiragana letter) matching. Because such lexical items are often presented in the form of kanji compounds, L2 learners need to be aware that kanji themselves have no phonological representations at the lexical level, and that it is important to know the exact sequence of hiragana letters if they would like to be a fluent writer with the keyboard. This learning strategy should also enhance reading comprehension skills, since learning dictionary representations entails an increase in the number of pronounceable kanji compound words. If they can go one step further to memorize the sound-meaning relation, it would become a perfect vocabulary building strategy because in this way, I believe, L2 vocabulary learning in Japanese will be able to lift itself up by its bootstraps.

How should L2 learners improve their visual recognition (reading) skills?

The answer to this question involves the issue of the role of reading, a discussion of the role of traditional rote writing practice, and a suggestion to apply L1 writing strategies to L2 writing instruction.

In a nutshell, learning compound kanji means learning lexical items, thus 'vocabulary growth' is the key word for this type of question. It would be possible for L2 kanji learners to say goodbye to their rather impractical orthographic-to-phonological computation strategy at a sub-lexical level, and gradually build up a 'feel' (i.e. a native-like intuitive judgment) for the lexicality of a given kanji compound word through extensive exposure to authentic printed Japanese, so that the amount of 'reading' should be another key to the improvement of L2 kanji recognition skills.

This native-like 'feel' can be fostered by widening individual knowledge of words to an idiomatic level of language use by becoming familiar with 'collocations' (see, (Sinclair, 1991)) in the larger context. In return, L2 learners can increase their semantic decoding power for unknown kanji words often found in the authentic printed materials. It is an obvious advantage for both language learners and teachers to have a sound knowledge of collocation in Japanese. Although there is no dynamic large corpus like the 'Bank of English' available in Japanese, some L2 and L1 Japanese researchers have started compiling smaller sized specialized spoken and written corpora (e.g. Kamada (1999), Yokoyama, Sasahara, Nozaki, and Long (1998)), so that in the near future, I hope, researchers, language teachers, and L2 learners will be able to gain frequency-based native intuitions such as 'collocable' lexical combinations (i.e. likely combinations of lexical items) through corpus search, without asking a native speaker of Japanese to give her or his native judgment.

Turning now to the writing strategy issue, as many L2 kanji teachers of Japanese actually say (Shimizu & Green, 2002) that traditional rote learning is their favorite instructional strategy, I may need to get back to the basics here with a suggestion that L2 learners should not underestimate the effectiveness of traditional kanji learning strategies such as rote writing by hand on long-term memory. This facilitative effect of kinetic memory (i.e. the effect of repeated writing) to recall visual kanji representations was empirically tested by L1 psycholinguistic studies such as Naka and Naoi (1995) and Nihei (1991), and evidence shows that this learning strategy is effective for free-recall rather than visual recognition. The L1 result implies that rote handwriting practice may be only effective to recall the visual representation of kanji for handwriting. This evidence may be also supported by an L1 writing strategy known as *kuusho* (i.e. writing kanji in the air or on the palm by the index-finger). Many Japanese develop a habit of *kuusho* for the purpose of an emergency aid to trace back the strokes of the intended kanji when the shape of that kanji has slipped their mind. Here, it is worth noting that the *kuusho* strategy is also widely used as an instructional method to teach primary school children in the L1 educational setting.

My suggestion is that this L1 *kuusho* strategy could be usable in an L2 setting when L2 writers look for visual cues to identify the intended kanji from a set of visually similar kanji compounds. As evidence of the L2 'kanji illusion' suggests, L2 readers' 'eyes' are rather unreliable under a time-pressured reading conditions, and that it would be better to have redundant multiple options than having to rely on a single information source. Moreover, the visual memory of lexical items must

constantly receive reinforcements by a method that is directly connected to the writer's long-term memory. At this point, I assume that rote handwriting practice, which requires short-term repetition of new kanji words until the writing reaches the autonomous level, as well as the related *kuusho* strategy, would serve best for this purpose, because of their direct motoric link to the long-term memory storage. Of course, we need further L2 research to test this idea, but if this L1 handwriting strategy can serve for L2 visual kanji discrimination purposes, it could bridge the gap between the radically different kanji formation processes of handwriting and digital writing.

Should L2 kanji learners still need to spend a lot of time practicing kanji by hand?

The answer to this question touches upon the use of electronic literacy tools in untutored naturalistic SLA by children and adults, and the future of kanji in the electronic age.

My answer is definitely 'yes' for the reason given above, but I would like to discuss this issue for those who are learning Japanese in a natural setting without attending any language courses, and those who learn Japanese as a second language during childhood. With the knowledge of the current computer-based writing situation in Japanese, it came as no surprise when I was first asked this question by an English person who had been living in Japan over a decade working as an English teacher. According to him, spending considerable time in remembering kanji using handwriting is just waste of his time, since he was certain that he could write reasonable Japanese with his PC; for example, if he came across an unknown kanji word, he could just plug in his on-line dictionary. So, what was the point of copying out kanji dozens of times every day? I just then wondered how many JFL (Japanese as a Foreign Language) teachers of Japanese had been prepared to answer this · radical question. Another British person who asked me the same question has a child whose mother is Japanese. His motivation was to relieve his child from sweating over getting his composition neat for kanji writing. Yet they have lived and will be living outside Japan, so that, according to the father, the kanji knowledge need not be as perfect as children living in Japan. Again the use of keyboard was mentioned. Although he was less radical than the English teacher above, they basically shared the same idea about rote handwriting kanji practice, and the possibility of deciding on the sole use of digital writing tools for their written communication.

Their key argument is that they are ready to opt for the acquisition of a 'reading only' lexicon, so that they would be free from burdensome kanji writing by hand, but that they do not need to compromise written communication for their ability to speak Japanese, which makes it possible to write using a keyboard if they can 'read' kanji. Obviously, it is beyond the scope of this thesis to consider what impact it would have on developing literacy to aim intentionally to acquire a recognition-only kanji lexicon. My answer to the above father was that the idea of encouraging the development of a reading-only lexicon was theoretically possible, though I had to say that Japanese people today still appreciate handwriting skills, for example, in their admiration of artistic brush lines of calligraphy, and one still needs a reasonable handwriting skill in the real-life situation in Japan. Nevertheless, if the parents concerned here would decide to educate their child in a novel way in order to cope with the L2 learning situation, we would witness a new type of literate person in the very near future. Ultimately, the very purpose of the human language, including orthography, is to communicate with other people, so if some people wish to achieve this goal in a foreign language by a rather unconventional way because of their need, why not support them?

Similarly in an L1 setting, Takata (1991) seems to have made some predictions about the effect of electronic literacy tools on the future of kanji, and predicts that there may develop two distinct lexicons of kanji; those that are used only for reading, and for writing using electronic tools, and a subset of these that are used only for handwriting. The truth hurts sometimes, but enthusiastic traditionalist kanji teachers must acknowledge that the current trend of literacy practice in L2 naturalistic settings has come remarkably far because of the new writing technology. At the same time, L2 literacy educators would probably need to keep their eye on the emergence of a similar trend in the JFL setting.

References

ACTFL (Ed.). (1999). *ACTFL:Oral Proficiency Interview: Tester Training Manual*. ACTFL, Inc.

Ball, E., & Blachman, B. (1988). Phoneme segmentation training: Effect on reading readiness. *Annals of Dyslexia, 38*, 208-225.

Barton, S., & Sanford, A. (1993). A case-study of anomaly detection: Shallow semantic processing and cohesion establishment. *Memory and Cognition, 21*, 477-487.

Bertoncini, J., Floccia, C., Nazzi, T., & Mehler, J. (1995). Morae and syllables: Rhythmical basis of speech representations in neonates. *Language and Speech, 38*(4), 311-329.

Bradley, D. C., Sanchez-Casas, R. M., & Gracia-Albea, J. E. (1993). The status of the syllable in the perception of Spanish and English. *Language and Cognitive Processes, 8(2)*, 197-233.

Bradley, L., & Bryant, P. (1985). *Rhyme and Reason in Reading and Spelling*. Ann Arbor, MI: The University of Michigan Press.

Bridwell-Bowles, L., Johnson, P., & Brehe, S. (1987). Composing and Computers: Case Studies of Experienced Writers. In A. Matsuhashi (Ed.), *Writing in Real Time: Modeling the Production Process*. Norwood, New Jersey: Ablex Publishing Corporation.

Bunkacho. (1971). *Gairai go no Meishi [The Noun Category in Foreign Loan Words], in Onsei to Onsei Kyooiku, Nihongo Kyooiku Shidoo Sankoosho vol.1*. Bunkacho (Agency for Cultural Affairs.

Bunkacho (Ed.). (1991). *Koyo bun no kaki arawashi kata no kijun [Criteria for writing government documents]*. Bunkacho (Agency for Cultural Affairs).

Burt, J. S., & Butterworth, P. (1996). Spelling in adults: Orthographic transparency, learning new letter strings and reading accuracy. *European Journal of Cognitive Psychology, 8*(1), 3-43.

Chen, M. J., & Yuen, J. C.-K. (1991). Effects of pinyin and script type on verbal processing: Comparisons of China, Taiwan, and Hong Kong experience. *International Journal of Behavioral Development, 14*(4), 429-448.

Chikamatsu, N. (1996). The effects of L1 orthography on L2 word recognition: A study of American and Chinese learners of Japanese. *Studies in Second Language Acquisition, 18*, 403-432.

Chung, S. (1995). *A Japanese Word Processor for Windows: JWP User's Mannual.*

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 100*(4), 589-608.

Cook, V. (1993). *Linguistics and Second Language Acquisition.* MACMILLAN.

Corder, S. (1973). *Introducing Applied Linguistics.* Penguin Books.

Corder, S. (1974). Error Analysis. In J. Allen & S. Corder (Eds.), *The edinburgh course in applied linguistics, vol. 3.* Oxford university Press.

Crystal, D. (1991). *A Dictionary of Linguistics and Phonetics [Third edition].* Basil Blackwell: Oxford.

Cunningham, A. E. (1990). Explicit versus implicit instruction in phonemic awareness. *Journal of Experimental Child Psychology*, 557-572.

Cunningham, A. E., & Stanovich, K. E. (1990). Assessing print exposure and orthographic processing skill in children: A quick measure of reading experience. *Journal of Educational Psychology, 82*(4), 733-740.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology, 24*, 381-410.

Cutler, A., & Otake, T. (1994). Mora or phoneme? Further evidence for language-specific listening. *Journal of Memory and Language, 33*, 824-844.

Cutler, A., & Young, D. (1994). Rhythmic structure of word blends in English. *Proceedings of ICSLP '94, Yokohama, Japan*, 1407-1410.

Dagneaux, E., Denness, S., & Granger, S. (1998). Computer-aided Error Analysis. *System, 26*, 163-174.

Dairoku, H. (1995). Mora ni taisuru ishiki wa kanamoji no yomi shutoku no hitsuyo joken ka? [Is awareness of morae a requisite for acquisition of kana reading?]. *Shinrigaku Kenkyu (The Japanese Journal of Psychology), 66*(4), 253-260.

Daiute, C. (1985). Do writers talk to themselves? In S. Freeman (Ed.), *The Acquisition of Written Language.* Norwood, New Jersey: Ablex Publishing Corporation.

DeFrancis, J. (1977). *Beginning Chinese Reader (2nd ed.).* New Haven.

Dowling, C. (1994). Word processing and the ongoing difficulty of writing. *Computers and Composition, 11*, 227-235.

Dunkel, P. (Ed.). (1991). *Computer-Assisted Language Learning and Testing: Research Issues and Practice.* Newbury House.

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: a perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance, 25(6)*, 1568-1578.

Ehri, L. C. (1980). The development of orthographic images. In U. Frith (Ed.), *Cognitive Processes in Spelling.* London: Academic Press.

Ehri, L. C. (1987). Learning to read and spell words. *Journal of Reading Behavior, 19*, 5-31.

Ehri, L. C. (1992). Reconceptualizing the development of sight word reading and its relationship to recoding [Part]. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading Acquisition* (pp. 107–143). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Ellis, N. C. (2002). Reflections on frequency effects in language processing. *Studies in Second Language Acquisition, 24*, 297-334.

Ellis, R. (1985). *Understanding Second Language Acquisition.* Oxford University Press.

Endo, O. (1989). Gairai go no Hyooki [Orthography of foreign loan words]. In *Kooza Nihongo to Nihongo Kyooiku: Moji Hyooki no Kyooiku, vol.8.* Tokyo: Meiji Shoin.

Erickson, T., & Mattson, M. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning and Verbal Behavior, 20*, 540-552.

Ericsson, K. A., & Simon, H. A. (1979). *Protocol Analysis: Verbal reports as data.* Cambridge, MA: MIT press.

Eysenck, M. W., & Keane, M. T. (2000). *Cognitive Psychology: A Student's Handbook.* East Sussex, Psychology Press Ltd.

Flaherty, M. (1993). Are Japanese kanji processed like pictures? *Psychologia, 36*, 144-150.

Flaherty, M., & Noguchi, M. S. (1998). Effectiveness of different approaches to kanji education with second language learners. *JALT Journal, 20*, 60-78.

Fleiss, J. L. (1981). *Statistical Methods for Rates and Proportions. 2nd ed.* New York: John Wiley.

Frith, U. (1980). *Cognitive Processes in Spelling.* London: Academic Press.

Gass, S. M., & Mackey, A. (2002). Frequency effects and second language acquisition. *Studies in Second Language Acquisition, 24*, 249-260.

Ghadessy, M., Henry, A., & Roseberry, R. L. (Eds.). (2001). *Small Corpus Studies and ELT: Theory and practice.* John Benjamins Publishing Company.

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 674-691.

Goswami, U., & Bryant, P. (1990). *Phonological Skills and Learning to Read.* Lawrence Erlbaum Associates Ltd, Publishers.

Goto, H., Fukasawa, N., & Hamada, M. (2001). The development of a computer textbook and its practical use. *Journal of Japanese Language Teaching, 110,* 150-159.

Gough, P. B., Ehri, L. C., & Treiman, R. (Eds.). (1992). *Reading Acquisition.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Granger, S. (1998). The Computer Learner Corpus: A versatile new source of data for SLA research. In S. Granger (Ed.), *Learner English on Computer.* Addison Wesley Longman.

Haas, C. (1996). *Writing Technology: Studies on the materiality of literacy.* Mahwah, NJ:Erlbaum.

Hadamitzky, W., & Spahn, M. (1997). *Kanji and Kana: Revised edition.* Tuttle.

Han, M. (1992). The timing control of geminate and single stop consonants in Japanese: A challenge for nonnative speakers. *Phonetica, 49,* 102-127.

Harada, E. (1999). Ninchi-teki Kadai Bunseki [Cognitive account of task analysis]. In H. Kaiho & T. Kato (Eds.), *Ninchi Kenkyu no Giho.* Fukumura Shuppan: Tokyo.

Haratsuchi, H. (1989). Kanji no shido ho: Kanji kei [How to teach kanji to speakers of logographic languages]. In A. Kato (Ed.), *Koza nihongo to nihongo kyoiku: Nihongo no moji hyoki (ge).* Tokyo: Meiji Shoin.

Harley, T. A. (1995). *The Psychology of Language: From Data to Theory.* Psychology Press.

Hatasa, K. (1989). *A study of learning and teaching kanji for non-native learners of Japanese.* Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Hatasa, K. (1997). Computer literacy for Japanese language professionals - what do we need to know about applications of the computer? In *Progress in japanese linguistics and pedagogy: A collection in honor of professor akira miura's 70th birthday.* ALC press.

Hatasa, Y. A. (2002). The effects of differential timing in the introduction of Japanese syllabaries on early second language development in Japanese. *The Modern Language Journal, 86(3),* 349-367.

Hatch, E., & Lazaraton, A. (1991). *The Research Manual: Design and Statistics for Applied Linguistics.* Newbury House Publishers.

Hatta, T., Kawakami, A., & Hatasa, Y. (1997). Kanji writing errors in Japanese college students and American Japanese students. In H.-C. Chen (Ed.), *Cognitive processing of chinese and related asian languages* (Vol. 74, p. 401-416). Hong Kong: The Chinese University Press.

Hatta, T., & Konda, T. (1992). Changes in visual field advantage in processing face and unknown letter with increasing stimulus familiarization. *Psychologia, 35,* 21-32.

Hayes, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry, 20,* 253-306.

Henderson, E. H. (1980). Developmental concept of word. In E. Henderson & J. Beers (Eds.), *Developmental and Cognitive Aspects of Learning to Spell: A reflection of word knowledge.* Newark, DE: International Reading Association.

Henderson, L. (1982). *Orthography and Word Recognition in Reading.* London: Academic Press.

Henshall, K. G., & Takagaki, T. (1990). *A guide to learning hiragana & katakana.* Tuttle.

Hill, C. A., Wallace, D. L., & Haas, C. (1991). Revising on-line: Computer technologies and the revising process. *Computers and Composition, 9(1),* 83-109.

Hirata, Y. (1999). *Acquisition of Japanese rhythm and pitch accent by English native speakers.* Unpublished doctoral dissertation, The University of Chicago.

Hirose, H. (1992). Jukugo no ninchikatei ni kansuru kenkyu - priming-ho ni yoru kento [Using the Priming Paradigm to Investigate Word Recognition for Kanji Compound Words]. *The Japanese Journal of Psychology, 63*(5), 303-309.

Horiba, Y. (1996). Comprehension process in L2 reading. *Studies in Second Language Acquisition, 18,* 433-473.

Horodeck, R. A. (1987). *The role of sound in reading and writing kanji.* Unpublished doctoral dissertation, Cornell University.

Inagaki, K., Hatano, G., & Otake, T. (2000). The effect of kana literacy acquisition on the speech segmentation unit used by Japanese young children. *Journal of Experimental Child Psychology, 75,* 70-91.

Inagaki, S. (1991). Gairaigo Hyooki no Kijyun to Kanyoo [The criteria and orthographic conventions for foreign loan words]. *Journal of Japanese Language Teaching, 74,* 71-80.

Ishida, T. (1989). Kanji no shido ho: Hi-kanji kei [How to teach kanji to speakers of non-logographic languages]. In A. Kato (Ed.), *Koza nihongo to nihongo kyoiku: Nihongo no moji hyoki (ge).* Tokyo: Meiji Shoin.

Ito, Y. (1989). Kanji no on kun [On and kun readings of kanji]. In Y. Takebe (Ed.), *Koza nihongo to nihongo kyoiku: Nihongo no moji hyoki (jo)*. Tokyo: Meiji Shoin.

Jim, M.-Z., Kabashima, T., & Murakami, M. (1993). Tegaki to wapro to ni yoru bunsho no keiryo bunseki [Quantitative analysis of sentences written by hand and word proccessor]. *Keiryo Kokugogaku (Mathematical Linguistics)*, *19*(3), 133–145.

Jorden, E. H. (1986). On teaching Nihongo. *Japan Quarterly*, *33*(2), 139-147.

Jorden, E. H., & Chaplin, H. I. (1976). *Reading Japanese*. New haven, CT: Yale University Press.

Kaiho, H. (1979). Kanji joho shori kisei o megutte [Information proccssing for kanji]. *Keiryo Kokugogaku*, *11*, 331-340.

Kaiho, H., & Inukai, Y. (1982). Kyoiku Kanji no Gaikei Tokucho no Shinri-teki Bunseki [An analysis of Gestalt characteristics of 881 Japanese/Chinese kanji]. *Kyoiku Shinrigaku Kenkyu*, *53*, 257-260.

Kaiho, H., & Nomura, Y. (1983). *Kanji joho shori no shinrigaku [The psychology of kanji processing]*. Tokyo: Kyoiku Shuppan.

Kamada, O. (1999). *KY Corpus to Daini Gengo toshite no Nihongo no Kenkyuu. [The KYCorpus and the study of Second Acquisition of Japanese] in Daini Gengo toshite no Nihongo Shuutoku ni Kansuru Soogoo Kenkyuu [The study of Second Language Acquisition of Japanese]* (Tech. Rep.). (Kaken Kenkyuu Hookokusho 08308019 (Contact person: Hiroko Quackenbush))

Kamas, E., Reder, L., & Ayers, M. (1996). Partial matching in the Moses illusion: Response bias not sensitivity. *Memory and Cognition*, *24*, 687-699.

Kanda, Y. (1986). Nihongo to computer [Japanese language and computer]. *Gengo (Language)*. (This text was available at http://www.fujitsu.co.jp/people/kanda/txt/gengo.txt)

Kawaguchi, Y., Kano, C., & Sakai, J. (Eds.). (1995). *Kanji shido idea book [Kanji teaching idea book]*. Tokyo: Sotakusha.

Kearns, R. (1994). *Prelexical speech processing by mono-and bilinguals*. Unpublished doctoral dissertation, University of Cambridge.

Kellogg, R. T. (1994). *The Psychology of Writing*. New York: Oxford University Press.

Kess, J. F., & Miyamoto, T. (1999). *The Japanese Mental Lexicon: Psycholinguistic Studies of Kana and Kanji Processing*. John Benjamins Publishing Company.

Kimura, I. (1988). *Wapuro tettei nyumon [A complete introduction of word processing]*. Tokyo: Iwanami Shoten.

Kimura, M., Sakata, Y., Kubota, T., & Kawamoto, T. (Eds.). (1989). *Nihongo kyoju ho [Methodologies for teaching Japanese]*. Tokyo: Ofusha.

Kobayashi, M., Quakenbush, H., & Fukada, A. (1991). Gairaigo ni Mirareru Nihognoka Kisoku no Shuutoku: Eigo washa no Choosa ni Motozuite [The acquisition of naturalization rules in loan words: The analysis of data from English-speaking learners of Japanese]. *Journal of Japanese Language Teaching, 74*, 48-59.

Koda, K. (1987). Cognitive strategy transfer in second language reading. In *Research in reading in english as a second language*. Washington, DC: TESOL.

Koda, K. (1988). Cognitive process in second language reading: Transfer of L1 reading skills and strategies. *Second Language Research, 4*, 135-156.

Koda, K. (1989). Effects of L1 orthographic representation on L2 phonological coding strategies. *Journal of Psycholinguistic Research, 18*, 201-222.

Koda, K. (1990). The use of L1 reading strategies in L2 reading: Effects of L1 orthographic structures on L2 phonological recoding strategies. *Studies in Second Language Acquisition, 12*, 393-410.

Koda, K. (1995). Second language reading research: Implications for jsl pedagogy. *In S. Makino (Ed.) The proceedings of the Third Princeton Japanese Pedagogy Wordshop*.

Komori, S. (1998). 5.an example of analysis: Error analysis of 'te-form ' in L2 writing corpus (in Japanese). *The Proceedings of the 1998 Conference for JCSS (http://langue.hyper.chubu.ac.jp/komori/JCSS98sec5.html)*.

Kondo, M. (1998). Heavy syllables and writing errors. *The Proceedings of the 1998 Conference for Japanese Langauge Teachers in Higher Education 12th September 1998, London*, 30-41.

Koyanagi, K. (2001). The role of cognition in second language acquisition processes (in Japanese). *Journal of Japanese Language Teaching, 109*, 10-19.

Kubozono, H. (1989). The mora and syllable structure in Japanese: Evidence from speech errors. *Language and Speech, 32*(3), 249-278.

Kubozono, H. (1995). Perceptual evidence for the mora in japanese [Part]. In B. Connell & A. Arvaniti (Eds.), *Phonology and Phonetic Evidence: Papers in Laboratory Phonology iv* (p. 141-156). Cambridge: Cambridge University Press.

Kubozono, H. (1996). Speech segmentation and phonological structure. In T. Otake & A. Cutler (Eds.), *Phonological Structure and Language Processing: Cross-Linguistic studies*. Berlin, New York: Mouton de Gruyter.

Kubozono, H. (1998). *Onseigaku・Oninron (in japanese)*. Tokyo: Kuroshio Publishers.

Kubozono, H. (1999). Mora and syllable. In N. Tsujimura (Ed.), *The Handbook of Japanese Linguistics* (p. 31-61). Oxford: Blackwell.

Kubozono, H. (2001). *Mora and syllable in Japanese.* Kobe University research report.

Ladefoged, P. (1982). *A Course in Phonetics.* Harcourt Brace Jovanovich Publishers.

Lam, F. S., & Pennington, M. C. (1993). The computer vs. the pen: A comparative study of word processing in a Hong Kong secondary classroom. *Perspectives, 5,* 53-75.

Leong, C. K. (1991). From phonemic awareness to phonological processing to language access in children developing reading proficiency [Part]. In D. J. Sawyer & B. J. Fox (Eds.), *Phonological awareness in reading: The evolution of current perspectives* (pp. 217–254). New York: Springer-Verlag.

Leong, C. K., & Tamaoka, K. (1995). Use of phonological information in processing kanji and katakana by skilled and less skilled Japanese readers. *Reading and Writing: An Interdisciplinary Journal, 7,* 377-393.

Liberman, I. Y. (1971). Basic research in speech and lateralization of language: Some implications for reading disability. *Bulletin of the Orton Society, 21,* 71-87. (Offprint from Haskins Status Report on Speech Research, 24, 51-66)

Liberman, I. Y., Shankweiler, D., Fischer, F. W., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology, 18,* 201-212.

Mack, A., & Rock, I. (1998). *Inattentional Blindness.* Cambridge, MA: MIT Press.

Mann, V. A. (1986a). Distinguishing universal and language-dependent levels of speech perception: Evidence from Japanese listeners' perception of English "l" and "r". *Cognition, 24*(3), 169-196.

Mann, V. A. (1986b). Phonological awareness: The role of reading experience. *Cognition, 24,* 65-92.

Mann, V. A. (1991). Phonological awareness and early reading ability: One perspective [Part]. In D. J. Sawyer & B. J. Fox (Eds.), *Phonological awareness in reading: The evolution of current perspectives* (pp. 191–215). New York: Springer-Verlag.

Mann, V. A., Tobin, P., & Wilson, R. (1987). Measuring phonological awareness through the invented spelling of kindergarten children. *Merrill-Palmer Quarterly, 33*(3), 365-391.

Matsumoto-Sturt, Y. (2002a). *Towards the development of a spoken learner corpus of Japanese OPI.* (Presented at the conference of Edinburgh OPI Symposium 2002, August 28-29, 2002)

Matsumoto-Sturt, Y. (2002b). *The detection of L2 kanji errors in Japanese word processing.* (Presented at the conference of Second Language Research Forum (SLRF) 2002 in Toronto, October 3-6, 2002.)

Matsumoto-Sturt, Y. (2003). Analysis of orthographic errors in Japanese word processing by English learners of Japanese: Are they all 'homophone' kanji errors? (in Japanese). *Journal of Japanese Language Teaching, 118*, 17-26.

Matsunaga, S. (1995). *The role of phonological coding in reading kanji* (Tech. Rep. No. 6). Second Language Teaching & Curriculum Center, University of Hawaii: University of Hawaii at Manoa.

Mehler, J., Dommergues, J.-Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior, 20*, 298-305.

Mori, K., & Yagihashi, T. (1989). *Nihongo wapuro no tanjo [The creation of a Japanese word processor].* Tokyo: Maruzen.

Mori, Y. (1998). Effects of first language and phonological accessibility on kanji recognition. *The Modern Language Journal, 82*, 69-82.

Mori, Y. (2002). Individual differences in the integration of information from context and word parts in interpreting unknown kanji words. *Applied Psycholinguistics, 23*, 375-397.

Mori, Y. (2003). The roles of context and word morphology in learning new kanji words. *The Modern Language Journal, 87*, 404-420.

Morita, F. (1989). Gendai kanazukai. In Y. Takebe (Ed.), *Koza nihongo to nihongo kyoiku: Nihongo no moji hyoki (jo).* Tokyo: Meiji Shoin.

Morton, J., Sasanuma, S., Patterson, K., & Sakuma, N. (1992). The organization of the lexicon in Japanese: single and compound kanji. *British Journal of Psychology, 83*, 517-531.

Muraki, M., & Nakaoka, N. (1990). Hatsuon to sokuon-eigo chuugokugo washa no hatsuon [Part]. In *Kooza Nihongo to Nihongo Kyooiku.* Tokyo: Meiji Shoin.

Muter, V., & Snowling, M. (1997). Grammar and phonology predict spelling in middle childhood [Part]. In R. Treiman (Ed.), (pp. 407–425). Netherlands: Kluwer Academic Publishers.

Nagae, S. (1992). Shiya bunkatsu kenkyu kara mita kanji joho shori [Kanji information processing as manifested in studies of divided visual field]. *Shinrigaku Hyoron (Japanese Psychological Review), 35*(3), 269-292.

Nagai, K. (1997). Mora timing by British learners of Japanese [Part]. In J. Leather & A. James (Eds.), *New Sounds 97: Proceedings of the Third International Symposium on the Acquisition of Second-language Speech* (pp. 240–248). Klagenfurt: University of Klagenfurt.

Naka, M., & Naoi, H. (1995). The Effect of Repeated Writing on Memory. *Memory and Cognition, 23*, 201-212.

Nelson, H. E. (1980). Analysis of spelling errors in normal and dyslexic children. In U. Frith (Ed.), *Cognitive Processes in Spelling*. London: Academic Press.

Neufeld, G. (1980). On the adult's ability to acquire phonology. *TESOL Quarterly, 14*, 285-298.

Neufeld, G. (1988). Phonological asymmetry in second-language learning and performance. *Language Learning, 38*, 531-560.

Nihei, Y. (1991). Bonyo Undo Puroguramu wo Meguru Undo Kioku Kenkyu no Tenkai: Shoji no Mondai wo Chushin ni [A Generalized Movement Program for Kinetic Memory: Focusing on the problem of handwriting]. *Shinrigaku Hyoron, 34*, 358-382.

Nikkei PC21. (1998). *konpon kara manabu pasokon kooza 3: dezital bunsho hen [Learning to use the PC from the basics 3: Digital text]*. Nikkei BP: Tokyo.

Nomura, Y. (1981). Kanji, kana hyokigo no joho shori - yomi ni oyobosu data suishin-kei shori to gainen-suishinkei shori no koka [The information processing of kanji, kana script: The effects of data-driven and conceptually-driven processing on reading]. *The Japanese Journal of Psychology, 51*(6), 327-334.

Norman, D., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In R. Davidson, G. Schwartz, & D. Shapiro (Eds.), *The Design of Everyday Things*. New York: Doubleday.

Ogawa, Y., Hayashi, D., et al. (Eds.). (1982). *Nihongo kyoiku jiten [An encyclopedia of Japanese language education]*. Tokyo: Taishukan Shoten.

Oguma, R. (2001). Acquisition of Japanese long vowels production. *Journal of Japanese Language Teaching, 109*, 110-117.

Oostendorp, H. van, & Kok, I. (1990). Failing to notice errors in sentences. *Language and Cognitive Processes, 5(2)*, 105-113.

Osumi, N. (Ed.). (1992). *Sho/Chu Gakko no Nihongo Nyuryoku [Japanese word processing at elementary and junior-high schools]*. Tokyo: Kyoiku Kaihatsu Kenkyujo.

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? speech segmentation in Japanese. *Journal of Memory and Language, 32*, 258-278.

Otake, T., Hatano, G., & Yoneyama, K. (1996). Speech segmentation by Japanese listeners. In T. Otake & A. Cutler (Eds.), *Phonological Structure and Language Processing: Cross-Linguistic Studies*. Berlin, New York: Mouton de Gruyter.

Otake, Y. (1987). *Japanese word processing as an instructional activity for university students learning Japanese as a foreign language.* Unpublished doctoral dissertation, The University of Texas at Austin.

Pennington, M. (1996). *The Computer and the Non-native Writer: A natural partnership.* Cresskill, NJ: Hampton Press.

Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading Acquisition.* Hillsdale: New Jersey: Lawrence Erlbaum Associates, Publishers.

Perfetti, C. A., & Zhang, S. (1991). Phonological processes in reading Chinese characters. *Journal of Experimental Psychology, 17,* 633-643.

Quackenbush, H., & Oso, M. (1990). *Gairaigo no Keisei to sono Kyooiku [The formation of loan words and their teaching]* (Tech. Rep.). National language Research Institute.

Quackenbush, H. C., & Toda, T. (1992). *English loan-words in Japanese - Acquisition of phonological rules by Australian students.* (Paper presented at the Second language Acquisition Forum of the Pacific 92, Sydney University)

Read, C. (1981). *Children's Creating Spelling.* London: Routledge and Kegan Paul.

Reder, L., & Kusbit, G. (1991). Locus of the Moses illusion: Imperfect encoding, retrieval, or match? *Journal of Memory and Language, 30,* 385-406.

Rensink, R., O'Regan, J., & Clark, J. (1997). To see nor not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8,* 368-373.

Roach, P. (1983). *English Phonetics and Phonology:A Practical Course.* Cambridge University Press.

Rumelhart, D., & Norman, D. (1982). Simulating a skilled typist: A study of skilled cognitive-motor performance. *Cognitive Science, 6*(2), 1-36.

Saito, H. (1981). Kanji to kana no yomi ni okeru keitaiteki fugoka oyobi oninteki fugoka no kento [Use of graphemic and phonemic encoding in reading kanji and kana]. *Shinrigaku Kenkyu*(52), 266-272.

Sakade, F. (1959). *A guide to reading & writing Japanese.* Tuttle.

Sakamoto, T. (1993). Eigo washa ni okeru 'te-kei ' keisei kisoku no shuutoku ni tsuite. *Journal of Japanese Language Teaching, 80,* 125-135.

Sakuma, N., Sasanuma, S., Tatsumi, I. F., & Masaki, S. (1998). Orthography and phonology in reading Japanese kanji words: Evidence from the semantic decision task with homophones. *Memory and Cognition, 26*(1), 75-87.

Sanders, B. (1994). *A is for ox: The collapse of literacy and the rise of violence in an electronic age.* New York:Vintage.

Sanford, A. (2002). Context, attention and depth of processing during interpretation. *Mind & Language, 17,* 188-204.

Sasanuma, S. (1975). Kana and kanji processing in Japanese aphasics. *Brain and Language, 2*, 369-383.

Sato, F., Sugiyama, N., Furuhata, M., & Kobayashi, Y. (Eds.). (1996). *Me kara uroko no kanji no hon [A kanji book that makes you feel as if the scales had fallen from your eyes]*. Tokyo: Takarajima Sha.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed model of word recognition and naming. *Psychological Review, 96*, 523-568.

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior, 23*, 383-404.

Sewell, D. (1990). *New Tools for New Minds: A Cognitive Perspective on the Use of Computers with Young Children*. London: Harvester Wheatsheaf.

Shelton, J., & Martin, R. (1992). How semantic is automatic semantic priming? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 1191-1209.

Shibatani, M. (1990). *The languages of Japan*. Cambridge: Cambridge University Press.

Shiffrin, R., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 127-190.

Shimamura, A. P. (1987). Word comprehension and naming: An analysis of English and Japanese orthographies. *American Journal of Psychology, 100*(1), 15-40.

Shimizu, H., & Green, K. E. (2002). Japanese language educators' strategies for and attitudes toward teaching kanji. *The Modern Language Journal, 86*, 227-241.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception, 28(9)*, 1059-1074.

Simons, D. J., & Levin, D. T. (1997). Change Blindness. *Trends in Cognitive Sciences, 1(7)*, 261-267.

Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature, 388*, 381-382.

Stanovich, K. E. (1988). The right and wrong places to look for the cognitive locus of reading disability. *Annals of Dyslexia, 38*, 154-177.

Sugito, M. (1989). Onsetsu ka haku ka- chooon, hatsuon, sokuon [Part]. In *Kooza Nihongo to Nihongo Kyooiku*. Tokyo: Meiji Shoin.

Sugiura, M., Oso, M., Ichikawa, Y., Okumura, M., Komori, S., Shirai, H., Takizawa, N., & Sotoike, T. (1997). Nihongo Gakushuusha no Sakubun Corpus:

Denshika ni yoru kyooyuu shigenka [A corpus of learner compositions: toward compilation of the digitized data for general use]. *Proceedings of The third annual conference of the Association for Natural Language Processing.*

Sukegawa, Y. (1993). Bogo betsu ni mita hatsuon no keikoo- ankeeto choosa no kekka kara [Part]. In *Nihongo onsei to nihongo kyooiku* (pp. 187–222). (Offprint from Ministry of Education Research Report [Monbushoo juuten ryooiki kenkyuu kenkyuu seika hookokusho])

Susser, B. (1998). The mysterious disappearance of word processing. *Computers and Composition, 15,* 347-371.

Takata, T. (1991). Kanji no unmei [The fate of kanji]. In *Gengo vol. 20* (p. 52-58). Tokyo: Taishuukan.

Takebe, Y. (1989). *Kanji no oshiekata: Nihongo wo manabu hi-kanji-kei gaikokujin no tame ni [How to teach kanji - for speakers of non-logographic language].* Tokyo: Aruku Ltd.

Takebe, Y. (1991). *Moji hyoki to nihongo kyoiku [Orthography and Japanese language education].* Bonjinsha.

Tamamura, F. (1991). Nihongo ni okeru Gari Yooso to Gairaigo [Foreign elements and loan words in Japanese]. *Journal of Japanese Language Teaching, 74,* 13-27.

Tamaoka, K., & Menzel, B. (1994). Nihongo kyoiku ni okeru romaji shiyo hihan no ronriteki kankyo ni kansuru gengoshinrigakuteki kosatsu [To use or not to use romaji - a psycholinguistic evaluation of the arguments against the use of romaji for teaching Japanese with learners of alphabetic mother tongues]. *Dokusho Kagaku (The Science of Reading), 38,* 104-116.

Tamaoka, K., & Takahashi, N. (1999). Kanji niji jukugo no shoji kodo ni okeru goishiyohindo oyobi shojiteki fukuzatsusei no eikyo [The effects of word frequency and orthographic complexity on the writing process of Japanese two-morpheme compound]. *The Japanese Journal of Psychology, 70*(1), 45-50.

Tarone, E. (2002). Frequency effects, noticing, and creativity. *Studies in Second Language Acquisition, 24,* 287-296.

Tateoka, Y. (2001). Questions and problem-solving strategies in the reading process (in Japanese). *Journal of Japanese Language Teaching, 111,* 66-75.

Taylor, I., & Taylor, M. M. (1995). *Writing and Literacy in Chinese, Korean and Japanese.* Amsterdam: John Benjamins Publishing Company.

Templeton, S., & Bear, D. R. (1992). A summary and synthesis: Teaching the lexicon to read and spell. In S. Templeton & D. R. Bear (Eds.), *Development of Orthographic Knowledge and the Foundations of Literacy: A memorial Festschrift for Edmund H. Henderson*. Hillsdale: New Jersey: Lawrence Erlbaum Associates, Publishers.

Toda, T. (1997). Mora rizumu shuutoku ni okeru sutoratejii to nihongo onsei no shidoo [Strategies for producing mora timing and the implications to instructions of Japanese phonetics] [Part]. In *Proceedings of the Society for Teaching Japanese as a Foreign Language* (pp. 191–196). Klagenfurt: University of Klagenfurt.

Toda, T. (1998). Mora to chuukan gengo no onsetsu koozoo [The mora and Interlanguage syllable structure] [Part]. In *Gaikokujin gakushuusha ni yoru nihongo onsei no shuutoku kenkyuu* (pp. 25–44). (Offprint from Tsukuba University Research Report)

Toda, T. (1999). *Usage and acquisition of accent of English loanwords by Japanese learners - an analysis based on conversational data from English, Chinese and Korean speakers* (Tech. Rep. No. Gengogaku-kei Kiyoo Bungei Gengo Kenkyuu: Gengo 36). University of Tsukuba.

Tollini, A. (1994). Kanji shidoo joo no jikei no jyuuyoo sa [The importance of form in the teaching of kanji]. *Japanese-Language Education around the Globe, The Japan Foundation Japanese Language Institue, vol. 4*, 107-116.

Tomita, T., & Sanada, K. (1991). *kyoshi yo nihongo handbook 2: Hyoki kaitei ban [Teachers' handbook for teaching Japanese language 2-orthography, New edition]*. The Japan Foundation Japanese Language Institute.

Treiman, R. (1985). Phonemic awareness and spelling: Children's judgments do not always agree with adults'. *Journal of Experimental Child Psychology, 39*, 182-201.

Treiman, R. (1986). The division between onsets and rimes in English syllables. *Journal of Memory and Language, 25*, 476-491.

Treiman, R. (1991). Children's spelling errors on syllable-initial consonant clusters. *Journal of Educational Psychology, 83*(3), 346-360.

Treiman, R. (1992). The role of intrasyllabic units in learning to read and spell [Part]. In L. C. E. Philip B. Gough & R. Treiman (Eds.), *Reading Acquisition* (pp. 65–106). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Treiman, R. (1993). *Beginning to Spell*. New York: Oxford University Press.

Treiman, R., Cassar, M., & Zukowski, A. (1994). What types of linguistic information do children use in spelling? The case of flaps. *Child Development, 65,* 1318-1337.

Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General, 124* (2), 107-136.

Treiman, R., & Zukowski, A. (1988). Units in reading and spelling. *Journal of Memory and Language, 27,* 466-477.

Tsuchida, A. (2001). Japanese vowel devoicing: Cases of consecutive devoicing environments. *Journal of East Asian Linguistics, 10(3),* 225-245.

Tsuchiya, J. (2001). *Gaikoku-jin no tame no nihongo keyboard nyuuryoku shien system no bogobetsu customize [Japanese EMI support system for non-native speakers, customized according to native language]* (Tech. Rep. No. 15: 373-380). Denki tsuushin fukyuu zaidan kenkyuu choosa hookokusho.

Uchida, T. (1993). Chuugokujin nihongo gakushuusha ni okeru chooon to sokuon no chookaku-teki ninchi no tokuchoo. *Kyooiku Shinrigaku Kenkyuu [The Japanese Journal of Educational Psychology], 41,* 414-423.

Ujihira, A., & Kubozono, H. (1994). A phonetic and phonological analysis of stuttering in Japanese [Part]. In *Proceedings of the third international conference on spoken language processing* (Vol. 3, pp. 1195–1199).

Umemura, M., Ichikawa, H., Honda, K., Toda, T., Harada, H., Katsuura, T., & Kikuchi, Y. (1995). Characteristics of roman character input in Japanese word processing on keystroke interval time. *Journal of Human Ergology, 24,* 171-181.

Underwood, J., & Underwood, G. (1990). *Computers and Learning.* London: Basil Blackwell Ltd.

Usami, Y. (2001). *Nihongo Gakushuusha ni yoru Nihongo Sakubun to Sono Bo-goyaku to no Taiyaku database, ver.2 [Database of compositions written by learners of Japanese, and their translated versions: ver:2.]* (Tech. Rep. No. CD-Rom version). National language Research Institute.

Van Orden, G. C. (1987). A Rows is a Rose: Spelling, sound, and reading. *Memory and Cognition, 15,* 181-198.

Wade-Woolley, L., & Siegel, L. S. (1997). The spelling performance of ESL and native speakers of english as a function of reading skill [Part]. In R. Treiman (Ed.), (pp. 387–406). Netherlands: Kluwer Academic Publishers.

Weekes, B., & Coltheart, M. (1996). Surface dyslexia and surface dysgraphia: Treatment studies and their theoretical implications. *Cognitive Neuropsychology, 13,* 277-315.

Winer, B. (1971). *Statistical principles in experimental design (2nd ed)*. New York: McGraw-Hill.

Wydell, T. N., Butterworth, B., & Patterson, K. (1995). The inconsistency of consistency effects in reading: The case of Japanese kanji. *Journal of Experimental Psychology, 21* (5), 1155-1168.

Wydell, T. N., Patterson, E. K., & Humphreys, W. G. (1993). Phonologically mediated access to meaning for kanji: Is a Rows still a Rose in Japanese kanji? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 491-514.

Yokoyama, S., Sasahara, H., Nozaki, H., & Long, E. (1998). *Shinbun Denshi Media no Kanji: Asahi shinbun CD-ROM ni yoru kanji hindohyo [Kanji in the electronic newspaper media: Kanji frequency tables from the Asahi newspaper CD-ROM], National Language Research Institute Project, Special Publication 1*. Tokyo: Senseido.

Kana Table

H	K	R	H	K	R	H	K	R	H	K	R	H	K	R
あ	ア	a	い	イ	i	う	ウ	u	え	エ	e	お	オ	o
か	カ	ka	き	キ	ki	く	ク	ku	け	ケ	ke	こ	コ	ko
さ	サ	sa	し	シ	shi	す	ス	su	せ	セ	se	そ	ソ	so
た	タ	ta	ち	チ	chi	つ	ツ	tsu	て	テ	te	と	ト	to
な	ナ	na	に	ニ	ni	ぬ	ヌ	nu	ね	ネ	ne	の	ノ	no
は	ハ	ha	ひ	ヒ	hi	ふ	フ	fu	へ	ヘ	he	ほ	ホ	ho
ま	マ	ma	み	ミ	mi	む	ム	mu	め	メ	me	も	モ	mo
や	ヤ	ya				ゆ	ユ	yu				よ	ヨ	yo
ら	ラ	ra	り	リ	ri	る	ル	ru	れ	レ	re	ろ	ロ	ro
わ	ワ	wa										を	ヲ	o
ん	ン	n												

H	K	R	H	K	R	H	K	R	H	K	R	H	K	R
が	ガ	ga	ぎ	ギ	gi	ぐ	グ	gu	げ	ゲ	ge	ご	ゴ	go
ざ	ザ	za	じ	ジ	ji	ず	ズ	zu	ぜ	ゼ	ze	ぞ	ゾ	zo
だ	ダ	da	ぢ	ヂ	ji	づ	ヅ	zu	で	デ	de	ど	ド	do
ば	バ	ba	び	ビ	bi	ぶ	ブ	bu	べ	ベ	be	ぼ	ボ	bo
ぱ	パ	pa	ぴ	ピ	pi	ぷ	プ	pu	ぺ	ペ	pe	ぽ	ポ	po

H	K	R	H	K	R	H	K	R
きゃ	キャ	kya	きゅ	キュ	kyu	きょ	キョ	kyo
しゃ	シャ	sha	しゅ	シュ	shu	しょ	ショ	sho
ちゃ	チャ	cha	ちゅ	チュ	chu	ちょ	チョ	cho
にゃ	ニャ	nya	にゅ	ニュ	hyu	にょ	ニョ	nyo
ひゃ	ヒャ	hya	ひゅ	ヒュ	hyu	ひょ	ヒョ	hyo
みゃ	ミャ	mya	みゅ	ミュ	my	みょ	ミョ	myo
りゃ	リャ	rya	りゅ	リュ	ryu	りょ	リョ	ryo
ぎゃ	ギャ	gya	ぎゅ	ギュ	gyu	ぎょ	ギョ	gyo
じゃ	ジャ	ja	じゅ	ジュ	ju	じょ	ジョ	jo
びゃ	ビャ	bya	びゅ	ビュ	byu	びょ	ビョ	byo
ぴゃ	ピャ	pya	ぴゅ	ピュ	pyu	ぴょ	ピョ	pyo

H: hiragana
K: katakana
R: romaji (Hepburn style)

The three romanization systems

Hepburn					Kunreishiki	Nihonshiki
a	i	u	e	o		
ka	ki	ku	ke	ko		
sa	shi	su	se	so	si	si
ta	chi	tsu	te	to	ti / tu	ti / tu
na	ni	nu	ne	no	hu	hu
ha	hi	fu	he	ho		
ma	mi	mu	me	mo		
ya		yu		yo		
ra	ri	ru	re	ro		
wa				o	o	wo
n						
ga	gi	gu	ge	go		
za	ji	zu	ze	zo	zi	di
da	ji	zu	de	do	zu	du
ba	bi	bu	be	bo		
pa	pi	pu	pe	po		
kya		kyu		kyo		
sha		shu		sho	sya / syu /syo	sya / syu /syo
cha		chu		cho	tya / tyu / tyo	tya / tyu / tyo
nya		hyu		nyo		
hya		hyu		hyo		
mya		my		myo		
rya		ryu		ryo		
gya		gyu		gyo		
ja		ju		jo	zya /zyu /zyo	dya /dyu / dyo
bya		byu		byo		
pya		pyu		pyo		

Romaji and Hiragana correspondence in typical IME

a

	あ	い	う	え	お		あ	い	う	え	お
	a	i	u	e	o		la	li	lu	le	lo
		yi	wu				xa	xi	xu	xe	xo
			whu					lyi		lye	
								xyi		xye	
								いぇ			
								ye			
							うぁ	うぃ		うぇ	うぉ
							wha	whi		whe	who
								wi		we	

k

	か	き	く	け	こ		きゃ	きぃ	きゅ	きぇ	きょ
	ka	ki	ku	ke	ko		kya	kyi	kyu	kye	kyo
	ca		cu		co						
			qu				くゃ		くゅ		くょ
	カ			ケ			qya		qyu		qyo
	lka			lke			くぁ	くぃ	くぅ	くぇ	くぉ
	xka			xke			qwa	qwi	qwu	qwe	qwo
							qa	qi		qe	qo
							kwa	qyi		qye	

g

	が	ぎ	ぐ	げ	ご		ぎゃ	ぎぃ	ぎゅ	ぎぇ	ぎょ
	ga	gi	gu	ge	go		gya	gyi	gyu	gye	gyo
							ぐぁ	ぐぃ	ぐぅ	ぐぇ	ぐぉ
							gwa	gwi	gwu	gwe	gwo

s

	さ	し	す	せ	そ		しゃ	しぃ	しゅ	しぇ	しょ
	sa	si	su	se	so		sya	syi	syu	sye	syo
		ci		ce			sha		shu	she	sho
		shi									
							すぁ	すぃ	すぅ	すぇ	すぉ
							swa	swi	swu	swe	swo

z

	ざ	じ	ず	ぜ	ぞ		じゃ	じぃ	じゅ	じぇ	じょ
	za	zi	zu	ze	zo		zya	zyi	zyu	zye	zyo
		ji					ja		ju	je	jo
							jya	jyi	jyu	jye	jyo

t

	た	ち	つ	て	と		ちゃ	ちぃ	ちゅ	ちぇ	ちょ
	ta	ti	tu	te	to		tya	tyi	tyu	tye	tyo
		chi	tsu				cha		chu	che	cho
			っ				cya	cyi	cyu	cye	cyo
			ltu				つぁ	つぃ		つぇ	つぉ
			xtu				tsa	tsi		tse	tso
			ltsu								

てゃ てぃ てゅ てぇ てょ

						tha とぁ	thi とぃ	thu とぅ	the とぇ	tho とぉ
d	だ da	ぢ di	づ du	で de	ど do	twa	twi	twu	twe	two
						ぢゃ dya	ぢぃ dyi	ぢゅ dyu	ぢぇ dye	ぢょ dyo
						でゃ dha	でぃ dhi	でゅ dhu	でぇ dhe	でょ dho
						どぁ dwa	どぃ dwi	どぅ dwu	どぇ dwe	どぉ dwo
n	な na	に ni	ぬ nu	ね ne	の no	にゃ nya	にぃ nyi	にゅ nyu	にぇ nye	にょ nyo
h	は ha	ひ hi	ふ hu fu	へ he	ほ ho	ひゃ hya	ひぃ hyi	ひゅ hyu	ひぇ hye	ひょ hyo
						ふゃ fya		ふゅ fyu		ふょ fyo
						ふぁ fwa fa	ふぃ fwi fi fyi	ふぅ fwu	ふぇ fwe fe fye	ふぉ fwo fo
b	ば ba	び bi	ぶ bu	べ be	ぼ bo	びゃ bya	びぃ byi	びゅ byu	びぇ bye	びょ byo
						ヴぁ va	ヴぃ vi	ヴ vu	ヴぇ ve	ヴぉ vo
						ヴゃ vya	ヴぃ vyi	ヴゅ vyu	ヴぇ vye	ヴょ vyo
p	ぱ pa	ぴ pi	ぷ pu	ぺ pe	ぽ po	ぴゃ pya	ぴぃ pyi	ぴゅ pyu	ぴぇ pye	ぴょ pyo
m	ま ma	み mi	む mu	め me	も mo	みゃ mya	みぃ myi	みゅ myu	みぇ mye	みょ myo
y	や ya		ゆ yu		よ yo	や lya xya		ゆ lyu xyu		よ lyo xyo
r	ら ra	り ri	る ru	れ re	ろ ro	りゃ rya	りぃ ryi	りゅ ryu	りぇ rye	りょ ryo
w	わ wa				を wo	**n** ん n nn				
	わ lwa xwa					n' xn				

Experiment 1: Hiragana copy followed by kana-kanji conversion task[1]

● 次のひらがな文をもう一度タイプして漢字かな混じり文に変換してください.

Re-type the following hiragana phrase or sentences, changing them into hiragana-kanji mixed text.

例　　しろい ようふく　　(*White clothes*)
　　　白い洋服

1.　うみ の さかな*　　　　　　　　(*A sea fish*)
2.　あかるい へや*　　　　　　　　(*A bright room*)
3.　ここ に おかね が あります　　(*There's some money here*)
4.　きっぷ を かった　　　　　　　(*I bought a ticket*)
5.　りょうり が すきだ　　　　　　(*I like cooking*)
6.　おかあさん と おとうさん*　　　(*Mum and dad*)
7.　けっこんしき に いった　　　　(*I went to a wedding*)
8.　おおきい くるま　　　　　　　(*A large car*)
9.　いぬ の びよういん　　　　　　(*A dog beauty salon*)
10.　ゆきぐに を たびする　　　　　(*To travel in the snow country*)
11.　おじいさん の むかしばなし　　(*Grandfather's stories about the old days.*)
12.　おにいさんのともだち　　　　　(*Elder brother's friend*)
13.　あおい そら に しろい くも*　　(*White clouds in the blue sky*)
14.　いっかげつ の よてい だった　(*It was planned for a month.*)
15.　しゃしん を とった　　　　　　(*I took a photo.*)
16.　くろい くるま の うんてんしゅ　(*The driver of a black car*)
17.　くらい もり　　　　　　　　　(*A dark forest*)
18.　みかづき の よる　　　　　　　(*The night of the new moon*)
19.　りんご と みかん*　　　　　　(*An apple and a tangerine*)
20.　しずかな きょうしつ　　　　　(*A quiet classroom*)
21.　まんいん でんしゃ　　　　　　(*A crowded train*)
22.　せかいじゅう の ひとびと　　　(*All the people in the world*)
23.　はるなつあきふゆ*　　　　　　(*Spring, summer, autumn ,winter*)
24.　りょこう に いった　　　　　　(*I went on a trip*)
25.　でんわ で はなした　　　　　　(*I talked on the phone*)
26.　えんそく の けいかく*　　　　(*The plan for the trip*)
27.　むずかしい じしょ　　　　　　(*A difficult dictionary*)
28.　うつくしい にほん の やま　　(*Beautiful Japanese mountains*)
29.　ちょうしょく を たべた　　　　(*I ate breakfast.*)
30.　あおい うわぎ の おんなのこ　(*The girl in the blue jacket*)

---

[1] Tasks, which are marked by asterisk, are taken from Osumi (1992: pp.19-20). Tasks were presented to the participants without English translations.

Experiment 2: Hiragana production task[1]

● 次の文章を読んで（　　）の中の英単語を漢字になおしてください．それをワープロで書いてください．

Read the following sentences once. Then replace English word in brackets with word(s) in kanji according to the instruction given.

Example: 日本の(mountain) は高いです。　　山

Practice: 朝から (rain) が降っています。

1. (gold) のネックレス。　　　　　　　　　　　　　(ki-N)
    (Gold) *necklace.*
2. 私は (English) の先生になりたいです。　　　　　(e-i-go)
    I want to be an (English) *teacher.*
3. 父は (bird) が好きです。　　　　　　　　　　　(to-ri)
    My father likes (birds).
4. (book) を読もうと思います。　　　　　　　　　(ho-N)
    I think that I will read (books).
5. 私は (watch) を買いました。　　　　　　　　　(to-ke-i)
    I have bought a (watch).
6. (one week) 間に一回は外食する。　　　　　　　(I-Q-shu-u)
    We eat out at least once (a week).
7. (Tokyo) は大きな都市です。　　　　　　　　　(to-u-kyo-u)
    (Tokyo) *is a big city.*
8. ナイフで指を (cut-past form)た。　　　　　　　(ki-Q-ta)
    I (cut) *my finger with a knife.*
9. (medicine) を飲んだほうがいいですよ。　　　　(ku-su-ri)
    You had better take some (medicine).
10. 暗いから (light) をつけてください。　　　　　(de-N-ki)
    It's dark, so would you please turn the (light) *on?*
11. 私は (fish) がきらいです。　　　　　　　　　(sa.ka.na)
    I don't like (fish).
12. あれは、あなたの (house) ですか。(No)、ちがいます。(i-e) (i-i-e)
    Is that your (house)?　　(No), *it's not.*
13. 今日、私の家に (come-te form)ください。　　　(ki-te)
    Would you please (come) *to my house today?*

---

[1] Tasks were presented without English translation. The spelling of the target words in brackets also did not appear in the participants' experimental material. N and Q mark moraic nasal and geminate consonants respectively. Long vowels /ee/ and /oo/, which have one-to-two correspondence between sound an hiragana letters, are written according to each hiragana sound in isolation as e-i or e-e sequence for /ee/ and o-o or o-u sequence for /oo/.

14. (face) を洗ってください。 (ka-o)
     Please wash your (face).

15. このバスは大学の前を(pass)ますか。 (to-o-ri)
     Is this bus going to (pass) *in front of the university?*

16. 日本語の (teacher) は三人います。 (se-N-se-i)
     There are three (teachers) *of Japanese.*

17. 友だちが遊びに(come-past form)。 (ki-ta)
     Friends (came) *to visit me.*

18. お(elder sister)さん はどこで働いていますか。 (ne-e)
     Where does your (elder sister) *work?*

19. テストの答えは先生に (ask-te from)ください。 (ki-i-te)
     Please (ask) *your teacher about answers to this test.*

20. 明日、手紙を (to write) つもりです。 (ka-ku)
     I intend (to write) *a letter tomorrow.*

21. 私たちの (one's life) で何人の友だちができますか。 (I-Q-sho-u)
     How many friends could we make in our (life-time)?

22. あなたの (hobby) は何ですか。 (shu-mi)
     What is your (hobby)?

23. (ice) は冷たい。 (ko-o-ri)
     (ice) *is cold.*

24. ここに (stamp) をはってください。 (kit-te)
     Please stick the (stamp) *here.*

25. あなたの (job) は何ですか。 (shi-go-to)
     What is your (job)?

26. あの人は (singer) です。 (ka-shu)
     That person is a (singer).

27. これから夕食を (to make) ます。 (tsu-ku-ri)
     I am going to (make) *supper now.*

28. 明日の (weather) をテレビで見ます。 (te-N-ki)
     I am going to watch tomorrow's (weather) *on TV.*

29. (together) にレストランに行きませんか。 (I-Q-sho)
     Shall we go to a restaurant (together)?

30. 明日、(furniture) を買うつもりです。 (ka-gu)
     I am planning to buy (furniture) *tomorrow.*

Experiment 3: Kanji error monitoring task

A: lists of kanji compounds[1]

<u>P+ targets</u>

1. この本の出帆社   *The <u>publisher</u> of this book.*
2. 勉強に盛を出す   *To put <u>a lot of work</u> into studying.*
3. 信仰沁を持つ   *To have <u>faith</u>.*
4. ショパンの代表昨は何ですか   *What is a <u>typical work</u> of Chopin?*
5. これは子共が好きなポケモンだ   *This is Pokemon, which <u>children</u> love.*
6. エベレスト登山を心見る   *To <u>attempt</u> to climb Everest.*
7. とても寒単な仕事   *A very <u>easy</u> job.*
8. 日本の蚊化を勉強する   *To study Japanese <u>culture</u>.*
9. 復雑なシステム   *A <u>complicated</u> system.*
10. 新発売の賞品   *A new <u>product</u>.*
11. イギリスの上琉階級   *The English <u>upper</u> class.*
12. 東京で日米首脳怪談があった   *There was a <u>summit</u> of J-A leaders in Tokyo.*
13. 三十年観も会っていない人   *Somebody who I haven't met <u>for</u> 30 years.*
14. 健康は弟一によく眠ることだ   <u>*The most*</u> *important thing for health is to sleep well.*
15. この問題を見討してください   *Please <u>investigate</u> this problem.*
16. ご中告ありがとう   *Thank you for the <u>advice</u>.*
17. 居間、現在の人々   *People of <u>today</u>.*
18. 今晩、七字に会いましょう   *Let's meet this evening at <u>seven</u>.*
19. 中ば無意識にテレビをつけた   *I turned on the TV <u>almost</u> without realizing.*
20. 終身経の囚人   <u>*A life*</u> *prisoner.*

<u>P- targets</u>

1. 夏の間ずっと暑かった   *It was hot <u>all</u> summer.*
2. 作家の仕事は諸説を書くこと   *The job of an author is to write <u>novels</u>.*
3. 一棚千円の雑誌   *A magazine which costs 1,000 yen per <u>copy</u>.*
4. 会議は牛後 5 時からです   *The meeting will start at 5<u>pm</u>.*
5. 手紙を焼やす   *To <u>burn</u> a letter.*
6. インド料理の材料   <u>*Ingredients*</u> *for Indian cooking.*
7. お巣子を売っている店   *A shop that sells <u>sweets</u>.*
8. 一庄に行きましょう   *Let's go <u>together</u>.*
9. 干真面目な人   *A <u>frivolous</u> person.*
10. かわいいペットへの愛増表現   <u>*Affection*</u> *towards a beloved pet.*
11. 結婚を洩仕込む   *To make a <u>proposal</u> of marriage.*
12. 息話まるような熱戦だった   *A <u>breathtakingly</u> close game.*

---

[1] Tasks were presented without English translation. Unmarked versions of error kanji were used in the experiment.

13. 追い風が強い   *The <u>tailwind</u> is strong.*

14. 色地階のセーター   *Two sweaters with <u>different</u> colours.*

15. 最級電車で帰る   *To go home by <u>the last</u> train.*

16. 鮭屋でワインを買う   *To buy wine at the <u>off-license</u>.*

17. 朝五時に起きたから寝い   *I'm <u>sleepy</u> because I got up at 5am.*

18. 新聞の体刊日   *A newspaper <u>holiday</u>.*

19. リルケは師事です   *Rilke is a <u>poet</u>.*

20. 和洋折哀の家   *A house in <u>both</u> western and Japanese styles.*

<u>Fillers</u>

1. やっと席が空いた   *At last I found a seat.*

2. これは彼が好きな犬だ   *This is the dog that he likes.*

3. フランスの国花は何ですか   *What is the national flower of France?*

4. この薬はよく効く   *This medicine works well.*

5. 熱いお茶が入りました   *Here is your hot tea.*

6. 料金の精算をする   *Pay the excess fare.*

7. 血行をよくする運動   *A sport that stimulates the circulation of the blood.*

8. 日本の夏は高温多湿だ   *Japanese summer is hot and humid.*

9. 事故に遭った   *I had an accident.*

10. 年の初めのあいさつ   *New Year's greeting.*

11. １０対６で試合に負けた   *We lost the game 10 to 6.*

12. 校則に従おう   *Let's follow the school rules.*

13. 大金を持っている人   *Somebody who has a lot of money.*

14. 三台の車に分かれて乗る   *We went separately in three cars.*

15. スポーツでいい汗かこう   *I'll have a good sweat by playing sports.*

16. 図書館で本を借りた   *I borrowed a book from the library.*

17. 漢字を覚える   *To remember kanji.*

18. 植木に水をやる   *To water the plants*

19. バスの乗客人数を確認する   *To check the number of passengers in the bus.*

20. 家族と旅に出た   *I set off on a trip with my family.*

B: Instructions and practice

● 次の文を読んでまちがって使われている漢字が入っているかどうかを見ただけで判断してくだ
さい（ワープロを使ってはいけません）。書かれている漢字が正しいと思ったら、文の右側に「正
しい」と書いてください。

Read the following sentences, and for each one, decide whether it includes wrong kanji. Make the decision only by looking, and without using the word processor. Write 'tadashii' (correct) if you judge that there is no mistake.

● 書かれている漢字がまちがっていると思う文には、その誤字を各文の下にコピーしてください。
それから、誤字を正しくした単語をコピーした字の下にワープロで出してください。

If you think a kanji has been written wrongly, make a copy of that kanji under the sentence, and then write your corrected version of the kanji underneath the copy.

● 正しい漢字が出せない時は「出せない」と書いてください。その時、思っていることを言葉にして
ください。例えば、「この漢字どっかで見たけど読めないなぁ」とか、「どうして出ないのかしら」
とか、「ほかの読み方で出してみようかな」など、思っていることをそのまま口にしてください。

If you can't type the correct kanji, type 'dasenai' (I cannot type it). When you do this, please say what you are thinking, (for example "I've seen this kanji somewhere before, but I can't read it").

それではいくつか練習してみましょう。*Now we will do some practice trials.*
練習１　二本人の友達

● この文は正しいと思う → 正しい　とタイプしてください。
If you think this sentence is correct, please type 'tadashii'(correct).

● この文には間違っている漢字があると思う。
If you think this sentence includes wrong kanji,

　　　　1) 漢字をコピーする　→　二
　　　　Copy the appropriate kanji.
　　　　2) 正しい漢字をタイプして出す→　日本人
　　　　Type the correct kanji.
　　　　3) 漢字が出せない時は　出せない　とタイプしてください。
　　　　If you can't type the correct kanji, type 'dasenai' (I cannot type it).

できたら次の問題に移ります。マウスを使って右一番下にある下向きの三角形をクリックすると、画
面が次のアイテムに変わります。

When you have finished, we will go on to the main task. Use the mouse, please click the double upside-down triangle at the bottom-right of the window to reach the items at on the

next screen.

練習2	北海道に始雪が降る	(*First snow falls in Hokkaido*)
練習3	湯名な音楽家	(*A famous musician*)
練習4	新しい自転車	(*A new bicycle*)
練習5	美しい貝岸線	(*A beautiful coastline*)

これで練習を終わります。*This is the end of the practice.*

- それでは、問題に入ります。*The task begins now.*
- 問題1が出たら、右下にある四角いボックスにある赤い丸を押してから問題を始めてください。
- 問題1は、右下の三角形を二回クリックすると出ます。

The first sentence can be reached by clicking the upside down double triangle on the bottom right of the window, twice with the mouse.

練習5　　北海道に始雪が降る

これで練習を終わります。This is the end of the practice.

● それでは、問題に入ります。The task begins now.

● 問題1が出たら、右下にある四角いボックスにある赤い丸を押してから問題を始めてください。

● 問題1は、右下の三角形を二回クリックすると出ます。The first sentence can be reached by clicking the upside down double triangle on the bottom right of the window, twice with the mouse.

THE UNIVERSITY OF EDINBURGH
Asian Studies – Japanese
Japanese 1: Writing class
Questionnaire

©YMS

This questionnaire will be used: (1) to assess your language background, and develop the most suitable syllabus for your needs, and (2) as a part of my research.   Please note that in the case of publication, your identity will never be disclosed. Please return this questionnaire to Yoko Matsumoto-Sturt by Monday, 22nd October (I will be seeing you a lot in Week 2!).

                           Family name   *First name*
Name (block capitals):_____   _____

Please read each question carefully and answer all the questions.

1. Are you a bilingual? (Please circle)   Yes   No

    ● If you circled *Yes*, please specify languages: _____,
       and then go to question 3 below.
    ● If you circled *No*, go to question 2.

2. What is your mother tongue (your first language)?   _____

    ● If it is Chinese, please give more details (e.g. Cantonese, Mandarin, etc):

    _____

3. Do you know any other language other than the above? (Please circle)   Yes   No

    ● If you circled *Yes*, please specify language(s): _____,
       and then go to question 4 below.
    ● If you circled *No*, go to question 4 below.

4. Have you been to Japan before? (Please circle)   Yes   No

    ● If you circled *Yes*, please answer questions a. to c. below, and then go to question 5:
    ● If you circled *No*, go to question 5 below.

    a. How long ago was your trip to Japan? _____ ago.
    b. How long did you stay in Japan? (e.g. days, weeks, years) _____
    c. How often do/did you go to Japan? (e.g. once, go to Japan every year, lived there age x to y):

PTO

5. Have you learned Japanese language before? (Please circle)   Yes   No

    - If you circled *Yes*, please identify your background knowledge by ticking the most appropriate information from a. to e. below, and then go to question 6.
    - If you circled *No*, this is the end of the questionnaire.

    a. Spoken Japanese only ___   b. Written Japanese (kana) only_____
    c. Written (kana and kanji) Japanese only _____
    d. Both spoken and written (hiragana and katakana) Japanese _____
    e. Both spoken and written (kana and kanji) Japanese ___

6. Where did you learn Japanese? (Please tick)
    a. In Japan __   b. Outside Japan __   c. Both in Japan and other place(s) __

    - If you ticked a. (In Japan), go to question 7 below.
    - If you ticked b. (Outside Japan), specify the place, and then go to question 8 below.
      Place (e.g. UK): _____
    - If you ticked c. (Japan + other places), go to question 9 below.

7. What is the highest level of education you have completed in Japan? (Please tick)
    a. Kindergarten _____   b. Primary school _____
    c. Secondary school up to some level ____   d. Completed secondary school ____
    e. Language course _____

8. How did you learn Japanese? (Please tick)
    a. Self study _____   b. Adult education _____
    c. GCSE ____   d. A Level _____
    e. Other (Please specify): _____

9. How did you learn Japanese? (Please describe):

    _____
    _____
    _____
    _____
    _____
    _____

Thank you very much for your time.
Yoko Matsumoto-Sturt

Moraic Awareness test material
*Reading: a period indicates a mora boundary and [ ] contains a bimoraic syllable

Condition	Num.	Item (kanji)	* Reading	Hiragana	Category	Class	Meaning
3S3M-V-	1	規則	ki.so.ku	きそく	SJ	N	regulation
3S3M-V-	2	口火	ku.ti.bi	くちび	Native	N	fuse
3S3M-V-	3	七時	si.ti.zi	しちじ	SJ	N	seven o'clock
3S3M-V-	4	寿司屋	su.si.ya	すしや	Native	N	sushi shop
3S3M-V-	5	明日	a.si.ta	あした	Native	N	tomorrow
3S3M-V-	6	目下	me.si.ta	めした	Native	N	one's junior
3S3M-V-	7	遅刻	ti.ko.ku	ちこく	SJ	N	being late
3S3M-V-	8	土筆	tu.ku.si	つくし	Native	N	horsetail
3S3M-V-	9	羊	hi.tu.zi	ひつじ	Native	N	sheep
3S3M-V-	10	不足	hu.so.ku	ふそく	SJ	N	insufficiency
3S3M-V+	11	田舎	i.na.ka	いなか	Native	N	countryside
3S3M-V+	12	売り場	u.ri.ba	うりば	Native	N	salesroom
3S3M-V+	13	男	o.to.ko	おとこ	Native	N	man
3S3M-V+	14	車	ku.ru.ma	くるま	Native	N	car
3S3M-V+	15	桜	sa.ku.ra	さくら	Native	N	cherry tree
3S3M-V+	16	持続	zi.zo.ku	じぞく	SJ	N	continuation
3S3M-V+	17	煙草	ta.ba.ko	たばこ	Loan	N	cigarette
3S3M-V+	18	花火	ha.na.bi	はなび	Native	N	fire work
3S3M-V+	19	昼寝	hi.ru.ne	ひるね	Native	N	a nap
3S3M-V+	20	予約	yo.ya.ku	よやく	SJ	N	reservation
4S4M-V-	21	北風	ki.ta.ka.ze	きたかぜ	Native	N	north wind
4S4M-V-	22	薬屋	ku.su.ri.ya	くすりや	Native	N	chemist
4S4M-V-	23	七月	si.ti.ga.tu	しちがつ	SJ	N	July
4S4M-V-	24	客室	kya.ku.si.tu	きゃくしつ	SJ	N	guest room
4S4M-V-	25	すき焼き	su.ki.ya.ki	すきやき	Native	N	sukiyaki
4S4M-V-	26	父の日	ti.ti.no.hi	ちちのひ	Native	N	father's day
4S4M-V-	27	疲れ目	tu.ka.re.me	つかれめ	Native	N	eye strain
4S4M-V-	28	挽肉	hi.ki.ni.ku	ひきにく	Mix	N	minced meat
4S4M-V-	29	複雑	hu.ku.za.tu	ふくざつ	SJ	N	complexity
4S4M-V-	30	克服	ko.ku.hu.ku	こくふく	SJ	N	overcome
4S4M-V+	31	雨水	a.ma.mi.zu	あまみず	Native	N	rain water
4S4M-V+	32	岩山	i.wa.ya.ma	いわやま	Native	N	rocky mountain
4S4M-V+	33	活躍	ka.tu.ya.ku	かつやく	SJ	N	activity
4S4M-V+	34	昨日	sa.ku.zi.tu	さくじつ	SJ	N	yesterday
4S4M-V+	35	竹の子	ta.ke.no.ko	たけのこ	Native	N	bamboo shoot
4S4M-V+	36	直立	tyo.ku.ri.tu	ちょくりつ	SJ	N	upright
4S4M-V+	37	毒薬	do.ku.ya.ku	どくやく	SJ	N	poison
4S4M-V+	38	友達	to.mo.da.ti	ともだち	Native	N	friend

4S4M-V+	39	発熱	ha.tu.ne.tu	はつねつ	SJ	N	fever
4S4M-V+	40	乗り物	no.ri.mo.no	のりもの	Native	N	vehicle
3S4M-N	41	一年	i.ti.[ne.n]	いちねん	SJ	N	one year
3S4M-N	42	富士山	hu.zi.[sa.n]	ふじさん	Mix	N	Mt .Fuji
3S4M-N	43	朝晩	a.sa.[ba.n]	あさばん	Mix	N	morning and evening
3S4M-N	44	国民	ko.ku.[mi.n]	こくみん	SJ	N	nation
3S4M-N	45	連絡	[re.n] ra.ku	れんらく	SJ	N	contact
3S4M-N	46	文学	[bu.n] ga.ku	ぶんがく	SJ	N	literature
3S4M-N	47	天ぷら	[te.n] pu.ra	てんぷら	Loan	N	tempura
3S4M-N	48	洗濯	[se.n] ta.ku	せんたく	SJ	N	laundry
3S4M-N	49	親切	[si.n] se.tu	しんせつ	SJ	N	kindness
3S4M-N	50	叔母ちゃん	o.ba.[cha.n]	おばちゃん	Native	N	auntie
3S4M-Q	51	錯覚	[sa.k.] ka.ku	さっかく	SJ	N	illusion
3S4M-Q	52	はっきり	[ha.k.] ki.ri	はっきり	Native	AD	clearly
3S4M-Q	53	徳利	[to.k.]ku.ri	とっくり	Loan?	N	sake bottle
3S4M-Q	54	毒けし	[do.k.] ke.si	どっけし	Native	N	antidote
3S4M-Q	55	密告	[mi.k.] ko.ku	みっこく	SJ	N	tip-off
3S4M-Q	56	一足	[i.s.] so.ku	いっそく	SJ	N	a pair
3S4M-Q	57	はったり	[ha.t.] ta.ri	はったり	Native	N	bluff
3S4M-Q	58	説得	[se.t.] to.ku	せっとく	SJ	N	persuasion
3S4M-Q	59	御法度	go.[ha.t.] to	ごはっと	SJ	N	be banned
3S4M-Q	60	札幌	[sa.p.] po.ro	さっぽろ	Loan?	N	Sapporo
3S4M-R	61	婆様	[ba.a] sa.ma	ばあさま	Native	N	old woman
3S4M-R	62	おやまあ	o.ya.[ma.a]	おやまあ	Native	INT	Dear me!
3S4M-R	63	椎茸	[si.i.] ta.ke	しいたけ	Native	N	shiitake mushroom
3S4M-R	64	嬉しい	u.re.[si.i]	うれしい	Native	ADJ	be glad
3S4M-R	65	留学	[ryu.u.] ga.ku	りゅうがく	SJ	N	study abroad
3S4M-R	66	朝夕	a.sa.[yu.u]	あさゆう	Native	N	morning and evening
3S4M-R	67	生活	[se.e.] ka.tu	せいかつ	SJ	N	living
3S4M-R	68	説明	se.tu.[me.e]	せつめい	SJ	N	explanation
3S4M-R	69	法律	[ho.o.] ri.tu	ほうりつ	SJ	N	law
3S4M-R	70	武士道	bu.si.[do.o]	ぶしどう	SJ	N	the way of the samurai
3S4M-dip	71	迂回路	u.[ka.i.] ro	うかいろ	SJ	N	bypass
3S4M-dip	72	縞馬	si.[ma.u.] ma	しまうま	Native	N	zebra
3S4M-dip	73	日帰り	hi.[ga.e.] ri	ひがえり	Native	N	day trip
3S4M-dip	74	似顔絵	ni.[ga.o.] e	にがおえ	Native	N	portrait
3S4M-dip	75	矢面	[ya.o.] mo.te	やおもて	Native	N	bear the brunt of
3S4M-dip	76	軽石	ka.[ru.i.] si	かるいし	Native	N	float stone
3S4M-dip	77	血液	ke.[tu.e.] ki	けつえき	SJ	N	blood
3S4M-dip	78	魚の目	[u.o.] no.me	うおのめ	Native	N	corn
3S4M-dip	79	野苺	[no.i.] ti.go	のいちご	Native	N	wild strawberry
3S4M-dip	80	燃え殻	[mo.e.] ga.ra	もえがら	Native	N	cinder

THE UNIVERSITY OF EDINBURGH

School of Asian Studies-Japanese

©YMS

Sounds in Japanese Language

Instruction:

You will be asked for an intuitive judgment every time you hear a Japanese word. Please write down how many sound units or rhythmic beats you perceive in that word when you hear it. For instance, *Toyota*: 1,2,3, and *ikebana*: 1,2,3,4. But if you feel there are more or fewer rhythmic beats, that is also fine. The important thing is to trust your own judgment, because there are no right or wrong answers. You may also feel some words a little bit confusing: e.g. *banzai*. You may wonder whether there are 2 or 3 beats (or even more) in a word like *banzai*. But you are not allowed to write your answer in an indecisive way like "2 or 3". You are required to circle <u>one and only one</u> answer per word.

	Family name		*First name*
Name (block capitals):	_____		_____

Year (please tick):   1st year _____   2nd year _____   4th year _____   Other _____

Practice:

Judge how many rhythmic beats you perceive when you hear the following Japanese words. You will hear the word's number in English, followed by the word being pronounced three times in Japanese. While you are listening to the word being pronounced, you must circle the number that you perceive to correspond to the number of rhythmic units. Please circle only ONE number per word.

a) | 1 | 2 | 3 | 4 | 5 |

b) | 1 | 2 | 3 | 4 | 5 |

Listening (Sounds in Japanese Language):

Judge how many rhythmic beats you perceive when you hear the following Japanese words. You will hear the word's number in English, followed by the word being pronounced three times in Japanese. While you are listening to the word being pronounced, you must circle the number that you perceive to correspond to the number of rhythmic units. Please circle only ONE number per word.

1. | 1   2   3   4   5 |
2. | 1   2   3   4   5 |
3. | 1   2   3   4   5 |
4. | 1   2   3   4   5 |
5. | 1   2   3   4   5 |
6. | 1   2   3   4   5 |
7. | 1   2   3   4   5 |
8. | 1   2   3   4   5 |
9. | 1   2   3   4   5 |
10. | 1   2   3   4   5 |
11. | 1   2   3   4   5 |
12. | 1   2   3   4   5 |
13. | 1   2   3   4   5 |
14. | 1   2   3   4   5 |
15. | 1   2   3   4   5 |
16. | 1   2   3   4   5 |
17. | 1   2   3   4   5 |
18. | 1   2   3   4   5 |
19. | 1   2   3   4   5 |
20. | 1   2   3   4   5 |
21. | 1   2   3   4   5 |
22. | 1   2   3   4   5 |
23. | 1   2   3   4   5 |

Experimental stimuli used in the Experiment

A list of experimental materials

	items	*meaning*	*pronunciation*	*frequency*
1.	大阪	*Osaka*	o-o-sa-ka	(5)
2.	料理	*cooking*	ryo-o-ri	(12)
3.	教室	*classroom*	kyo-o-si-tu	(8)
4.	将来	*future*	sho-o-ra-i	(6)
5.	用事	*business*	yo-o-ji	(17)
6.	朝食	*breakfast*	cho-o-sho-ku	(4)
7.	公園	*park*	ko-o-e-n	(13)
8.	東京	*Tokyo*	to-o-kyo-o	(24)
9.	高速	*high speed*	ko-o-so-ku	(1)
10.	小説	*novel*	sho-o-se-tu	(2)
11.	妹	*younger sister*	i-mo-o-to	(15)
12.	飛行機	*airplane*	hi-ko-o-ki	(6)
13.	日曜日	*Sunday*	ni-ti-yo-o-bi	(24)
14.	冷蔵庫	*refrigerator*	re-e-zo-o-ko	(3)
15.	高校生	*high-school student*	ko-o-ko-o-se-e	(2)
16.	お正月	*New year*	o-sho-o-ga-tu	(4)
17.	お父さん	*father*	o-to-o-sa-n	(22)
18.	銀行員	*bank clerk*	gi-n-ko-o-in	(1)
19.	弟	*younger brother*	o-to-o-to	(12)
20.	自動車	*car*	ji-do-o-sha	(2)
21.	多少	*more or less*	ta-sho-o	(1)
22.	学校	*school*	ga-k-ko-o	(12)
23.	旅行	*traveling*	ryo-ko-o	(19)
24.	社長	*president*	sha-cho-o	(4)
25.	水道	*water supply*	su-i-do-o	(2)
26.	一生	*one's life*	i-s-sho-o	(1)
27.	勉強	*study*	be-n-kyo-o	(37)
28.	空港	*airport*	ku-u-ko-o	(2)
29.	食堂	*dining room*	sho-ku-do-o	(11)
30.	卒業	*graduation*	so-tu-gyo-o	(4)

3. 妹

これは＿＿＿＿と読みます。

Ranking of all the PRH correct words

R-1	日曜日	*Sunday*	ni-ti-yo-o-bi	(24)
R-2	勉強	*study*	be-n-kyo-o	(37)
R-2	お父さん	*father*	o-to-o-sa-n	(22)
R-4	東京	*Tokyo*	to-o-kyo-o	(24)
R-4	水道	*water supply*	su-i-do-o	(2)
R-6	料理	*cooking*	ryo-o-ri	(12)
R-6	社長	*president*	sha-cho-o	(4)
R-6	弟	*younger brother*	o-to-o-to	(12)
R-9	公園	*park*	ko-o-e-n	(13)
R-9	学校	*school*	ga-k-ko-o	(12)
R-11	教室	*classroom*	kyo-o-si-tu	(8)
R-11	自動車	*car*	ji-do-o-sha	(2)
R-13	食堂	*dining room*	sho-ku-do-o	(11)
R-14	用事	*business*	yo-o-ji	(17)
R-14	銀行員	*bank clerk*	gi-n-ko-o-in	(1)
R-14	高校生	*high-school student*	ko-o-ko-o-se-e	(2)
R-17	飛行機	*airplane*	hi-ko-o-ki	(6)
R-18	お正月	*New year*	o-sho-o-ga-tu	(4)
R-18	妹	*younger sister*	i-mo-o-to	(15)
R-18	大阪	*Osaka*	o-o*-sa-ka	(5)
R-18	将来	*future*	sho-o-ra-i	(6)
R-22	旅行	*traveling*	ryo-ko-o	(19)
R-22	朝食	*breakfast*	cho-o-sho-ku	(4)
R-24	多少	*more or less*	ta-sho-o	(1)
R-24	卒業	*graduation*	so-tu-gyo-o	(4)
R-24	空港	*airport*	ku-u-ko-o	(2)
R-27	小説	*novel*	sho-o-se-tu	(2)
R-28	高速	*high speed*	ko-o-so-ku	(1)
R-29	一生	*one's life*	i-s-sho-o	(1)
R-30	冷蔵庫	*refrigerator*	re-e-zo-o-ko	(3)

A list of kanji materials (target, non-words, homophone) and two types of sentence contexts

(c+ = highly related context, c- = less related context)

(v+ = visually similar, v- = visually dissimilar non-words)

TARGET (meaning)	V+	V-	HOMOPHONE (meaning)

1. 有効 (*effective*)   有絞   遊高   友好 (*friendly*)
/yu.u-ko.o/   /yu.u-ko.o/   /yu.u-ko.o/

> c+: 通勤中の姉は三年間も（　　　　　）な旅券を拾いました。
> *My sister picked up a passport, which is valid for three years, on her way to work.*
> c-: 仕事中の姉は集会場で（　　　　　）な物を貰いました。
> *My sister was given something valid while she was working.*

2. 解職 (*dismissal*)   解織   階植   海食 (*wave erosion*)
/ka.i-sjo.ku/   /ka.i-sho.ku/   /ka.i-sho.ku/

> c+: 十二月に（　　　）された市長は公金を悪用していた。
> *The mayor, who was dismissed in December, abused public money.*
> c-: 十二月に（　　　）された母親と旅行を計画している。
> *I am planning a trip with my mother who was dismissed in December.*

3. 終決 (*an end*)   終訣   州欠   集結 (*gather*)
/sju.u-ke.tu/   /sju.u-ke.tu/   /sju.u-ke.tu/

> c+: 一部の人は戦争の（　　　　　）という結末を期待している。
> *Some people are expecting the end of war as a final outcome.*
> c-: 一部の人は物事の（　　　　　）という意味を理解できない。
> *Some people cannot understand the meaning of the end of things.*

4. 人生 (*one's life*)   人性   陣聖   仁政 (*benevolent government*)
/ ji.N-se.e/   / ji.N-se.e/   / ji.N-se.e/

> c+: 数々の苦難を経験する（　　　　　）の物語を読んでいます。
> *I am reading a story about the life of somebody who experiences many hardships.*
> c-: 色々な名画を鑑賞する（　　　　　）の物語を読んでいます。
> *I am reading a story about the life of somebody who appreciates various works of art.*

TARGET (meaning)	V+	V-	HOMOPHONE (meaning)

5. 打倒 (*overthrouw*)   打到   駄投   妥当 (*appropriate*)
    /da-to.o/   /da-to.o/   /da-to.o/

> c+: 水泳の選手は海外の強敵「(   ) カナダ」が目標です。
> *The aim of a swimmer is to 'Beat Canada', which is a strong oversea opponent.*
> c-: 本店の店長は話題の新作「(   ) オペラ」が大好きだ。
> *The manager of the main branch loves 'Beat Opera', which is a famous new product.*

6. 近郊 (*the suburb*)   近効   禁口   金鉱 (*a gold mine*)
    /ki.N-ko.o /   /ki.N-ko.o /   /ki.N-ko.o /

> c+: 旅行案内に登場する京都の (   )の神社はきれいです。
> *The shrine in the Kyoto suburbs, which is featured in a travel guide, is beautiful.*
> c-: 空想小説に登場する未来の (   )の写真はきれいです。
> *The picture of the suburbs of the future, which is featured in a science fiction story, is beautiful.*

7. 時機 (*a chance*)   時幾   磁希   次期 (*a curtail time, a season*)
    /ji-ki /   /ji-ki /   /ji-ki /

> c+: 厳しい時代を必死で生きる人には (   ) がくるのが待ち遠しい。
> *People, who are surviving difficult times, can hardly wait for an opportunity to come.*
> c-: 暖かい居間で家族と楽しむ人には (   ) についての意識はない。
> *People, who are enjoying life with their family in a warm living room, do not notice opportunities.*

8. 長官 (*the Chief Cabinet Secretary*)   長管   町漢   朝刊 (*a morning edition paper*)
    /cjo.o-ka.N /   /cjo.o-ka.N /   /cjo.o-ka.N /

> c+: 軍隊の緊急事態に対応する国務 (   ) が電話に出た。
> *The Secretary of State, who has been dealing with Military emergency situations, answered the call.*
> c-: 去年の日本映画で主役だった山岸 (   ) の男優を見た。
> *I saw a leading actor who was featured in a last year's Japanese film as Mr Yamagashi, the Chief Cabinet Secretary.*

9. 漂泊 (*wandering*)   漂柏   評博   表白 (*expression*)
    /hjo.o-ha.ku /   /hjo.o-ha.ku /   /hjo.o-ha.ku /

> c+: 嵐の海上で何日も (   ) を続けた漁船を見つけた。
> *We found a fishing boat that had been drifting for many days in the stormy sea.*
> c-: 昔の地図で何回か (   ) を示した赤丸を見つけた。
> *We found red circles, which show a shipwreck, a couple of times in an old map.*

	TARGET (meaning)	V+	V-	HOMOPHONE (meaning)

10. 気候 (*climate*)   気喉   記広   帰港 (*return to port*)
/ki-ko.o /   /ki-ko.o /   /ki-ko.o /

c+: 最近、特に温暖になった(   ) について話しました。
We have recently talked about the climate that has become particularly warmer.
c-: 最近、特に話題になった(   ) について話しました。
We have recently talked about the climate that has been discussed.

11. 操作 (*operation*)   操詐   早左   捜査 (*criminal investigation*)
/so.o-sa/   /so.o-sa/   /so.o-sa/

c+: オペレーターは画面の上でマウスを（   ） しています。
The operator is controlling the mouse on the screen.
c-: マネージャーは机の上で光る物を（   ） しています。
The manager is operating a shiny thing on the desk.

12. 反抗 (*resist*)   反坑   半光   藩校 (*a feudal school*)
/ha.N-ko.o/   /ha.N-ko.o/   /ha.N-ko.o/

c+: 人種的な理由で （   ） した若者は後で有名になった。
A young man, who resisted against the race problem, became famous later on.
c-: 散歩中に道路で （   ） した子犬は後で病気になった。
A puppy, which disobeyed his owner while walking, became sick later on.

13. 保健 (*preservation of health*)   保鍵   歩県   保険 (*insurance*)
/ho-ke.N/   /ho-ke.N/   /ho-ke.N/

c+: ミネラル入り天然水を （   ） のために飲んでいる若者が多い。
There are many young people who drink natural mineral water to keep their health.
c-: ロンドンの動物学者が （   ） について書いた新書を見つけた。
I found a new book on health that was written by a zoologist in London.

14. 心境 (*a state of mind*)   心鏡   新胸   親共 (*offerings*)
/si.N-kjo.o /   /si.N-kjo.o /   /si.N-kjo.o /

c+: 試合で大負けして泣く水泳の選手の （   ） を聞くのは難しい。
It is difficult to ask about the feelings of a swimmer who has been beaten hollow at the competition.
c-: 学会で数学について話した大学の先生が （   ） を付け加えた。
A university lecturer, who talked about mathematics at the conference, added his feelings.

	TARGET (meaning)	V+	V−	HOMOPHONE (meaning)

15. 訪問 (*a visit*)   訪聞   方紋   法門 (*the gate to Buddhism*)
/ho.o-mo.N/   /ho.o-mo.N/   /ho.o-mo.N/

c+: 鈴木市長が姉妹都市カルガリの（   ）計画を出した。
The mayor Suzuki submitted a visiting plan to the sister city Calgary.
c−: 女性雑誌の写真広告にきれいな（   ）写真を見つけた。
I found beautiful photos of a visit in an advertisement in a woman's magazine.

16. 主張 (*assertion*)   主帳   酒聴   朱鳥 (*the name of an era : `shuchoo'*)
/sju-cjo.o /   /sju-cjo.o /   /sju-cjo.o /

c+: がんこな老人には曲げられない（   ）や態度が多い。
A stubborn old man often sticks with his position and does not change his opinion for anything.
c−: そこの雑誌に書いてあった（   ）や観察は古いです。
The opinions and observations in that magazine there are old.

17. 恐怖 (*fear*)   恐布   強符   教父 (*a godfather*)
/kjo.o-fu/   /kjo.o-fu/   /kjo.o-fu/

c+: ホンコンの高層ビルで次々に（   ）や不安の場面を撮った。
We shot the scenes of fear and anxiety one after another at a skyscraper in Hong Kong.
c−: フランスの画家ルネは次々に（   ）や愛情を題材に描いた。
The French painter Rene painted his pictures with a series of motifs on fear and love.

18. 業績 (*achievement*)   業積   凝石   行跡 (*conduct*)
/gjo.o-se.ki/   /gjo.o-se.ki/   /gjo.o-se.ki/

c+: 円高の影響で企業のこれまでの（   ）を見直した。
Companies reconsidered their current business results under the influence of the strong Yen.
c−: 米国の自宅の電話でこれからの（   ）を意見した。
I gave advice about the forthcoming business by telephone at home in the U.S.A.

19. 会議 (*meeting*)   会犠   回技   懐疑 (*doubt*)
/ka.i-gi/   /ka.i-gi/   /ka.i-gi/

c+: 来年、この会社で重要な（   ）を開催すると決まりました。
It has been decided that an important meeting will be held in this company next year.
c−: 去年、この地区で別々の（   ）を検討したと聞いています。
I have heard that there were separate meetings in this area last year.

	TARGET (meaning)	V+	V-	HOMOPHONE (meaning)

20. 幸福 (*happiness*)   幸幅   考服   降伏 (*surrender*)
/ko.o-fu.ku /   /ko.o-fu.ku /   /ko.o-fu.ku /

> c+: 姉の結婚後の生活は（　　　　　　　）になったと聞きました。
> *I have heard that my sister's life became happy after her marriage.*
> c-: 私の留学中の友達は（　　　　　　　）になったと聞きました。
> *I have heard that my friend, who has been studying abroad, became happy.*

21. 兄弟 (*brother*)   兄第   京台   強大 (*mighty*)
/kjo.o-da.i /   /kjo.o-da.i   /kjo.o-da.i

> c+: その男の子の家まで行くと（　　　　　　　）の一人が出てきました。
> *When I visited that boy's house, one of his brothers came out.*
> c-: その細い道を歩いて行くと（　　　　　　　）の山林が見えてきます。
> *If you walk along that narrow road, (my) brother's forest will gradually come into view.*

22. 急増 (*a rapid increase*)   急憎   級象   旧蔵 (*stored for a long time*)
/kju.u-zo.o/   /kju.u-zo.o/   /kju.u-zo.o/

> c+: 週末から降り続いた大雨で川の水が（　　　　　　）した。
> *The river has swollen rapidly because of the heavy rainfall since the weekend.*
> c-: 学校の窓から見える校庭で赤い色が（　　　　　　）した。
> *We can see from the school windows that there was a rapid increase in red colour in the schoolyard.*

23. 千倍 (*thousand times*)   千培   戦梅   専売 (*a monopoly*)
/se.N-ba.i /   /se.N-ba.i /   /se.N-ba.i /

> c+: 入社の競争率が（　　　　　）の会社でも働きたい気持ちがある。
> *I want to work for a company that has a one-in-a-thousand rate of competition for joining the company.*
> c-: 車内の運転手に（　　　　　）の気持ちで伝えたい用事がある。
> *I have something else to tell the driver in the car, with 1000 times more of my feelings.*

24. 公共 (*public*)   公供   校強   交響 (*symphony*)
/ko.o-kjo.o /   /ko.o-kjo.o /   /ko.o-kjo.o /

> c+: 夜間は病院などの（　　　　　　　）の場所では静かに話しましょう。
> *Let's talk quietly in a public place such as a hospital in the evening.*
> c-: 来年は仕事などで（　　　　　　　）を意識する人たちと話しましょう。
> *Next year, let's talk to people who have a public awareness about their work.*

Read the following sentences, and circle the right kanji in the list. You should circle only ONE kanji compound per question. Please also circle a number underneath each item, representing how certain you are about your choice. The numbers range from 1 to 5, where 1 represents completely unsure, and 5 represents completely certain, and the numbers in the middle represent degrees of certainty between these extremes.

名前 : _____

1. 通勤中の姉は三年間も_____な旅券を拾いました。
   a) 有絞   b) 遊高   c) 友好   d) 有効
   Completely unsure   1   2   3   4   5   *Completely certain*

2. 十二月に_____された市長は公金を悪用していた。
   a) 解職   b) 解織   c) 階植   d) 海食
   Completely unsure   1   2   3   4   5   *Completely certain*

3. 一部の人は戦争の_____という結末を期待している。
   a) 集結   b) 終訣   c) 終決   d) 州欠
   Completely unsure   1   2   3   4   5   *Completely certain*

4. 数々の苦難を経験する_____の物語を読んでいます。
   a) 人生   b) 人性   c) 陣聖   d) 仁政
   Completely unsure   1   2   3   4   5   *Completely certain*

5. 水泳の選手は海外の強敵「_____カナダ」が目標です。
   a) 打倒   b) 打到   c) 駄投   d) 妥当
   Completely unsure   1   2   3   4   5   *Completely certain*

6. 旅行案内に登場する京都の_____の神社はきれいです。
   a) 近効   b) 金鉱   c) 禁口   d) 近郊
   Completely unsure   1   2   3   4   5   *Completely certain*

7. 厳しい時代を必死で生きる人には＿＿＿＿＿＿＿＿がくるのが待ち遠しい。
　　a) 次期　　　　　　　b) 時機　　　　　c) 時幾　　　　　d) 磁希

Completely unsure　　1　　2　　3　　4　　5　*Completely certain*

8. 軍隊の緊急事態に対応する国務＿＿＿＿＿＿＿＿が電話に出た。
　　a) 朝刊　　　　　　　b) 長管　　　　　c) 長官　　　　　d) 町漢

Completely unsure　　1　　2　　3　　4　　5　*Completely certain*

9. 嵐の海上で何日も＿＿＿＿＿＿＿を続けた漁船を見つけた。
　　a) 評博　　　　　　　b) 漂泊　　　　　c) 漂柏　　　　　d) 表白

Completely unsure　　1　　2　　3　　4　　5　*Completely certain*

10. 最近、特に温暖になった＿＿＿＿＿＿＿について話しました。
　　a) 気候　　　　　　　b) 気喉　　　　　c) 記広　　　　　d) 帰港

Completely unsure　　1　　2　　3　　4　　5　*Completely certain*

11. オペレーターは画面の上でマウスを＿＿＿＿＿＿＿しています。
　　a) 操作　　　　　　　b) 早左　　　　　c) 操詐　　　　　d) 捜査

Completely unsure　　1　　2　　3　　4　　5　*Completely certain*

12. 人種的な理由で＿＿＿＿＿＿＿した若者は後で有名になった。
　　a) 半光　　　　　　　b) 反坑　　　　　c) 反抗　　　　　d) 藩校

Completely unsure　　1　　2　　3　　4　　5　*Completely certain*

13. ミネラル入り天然水を＿＿＿＿＿＿＿のために飲んでいる若者が多い。
　　a) 保険　　　　　　　b) 保健　　　　　c) 歩県　　　　　d) 保鍵

Completely unsure　　1　　2　　3　　4　　5　*Completely certain*

14. 試合で大負けして泣く水泳の選手の＿＿＿＿＿＿を聞くのは難しい。
　　a) 心鏡　　　b) 新胸　　　　　c) 親共　　　　d) 心境

Completely unsure　　1　　2　　3　　4　　5　*Completely certain*

15. 鈴木市長が姉妹都市カルガリの＿＿＿＿＿＿＿計画を出した。
　　a) 訪問　　　　　　　b) 訪聞　　　　　c) 方紋　　　　　d) 法門

Completely unsure　　1　　2　　3　　4　　5　*Completely certain*

16. がんこな老人には曲げられない＿＿＿＿＿＿＿＿＿や態度が多い。
    a) 主張   b) 主帳   c) 酒聴   d) 朱鳥
    Completely unsure   1   2   3   4   5   *Completely certain*

17. ホンコンの高層ビルで次々に＿＿＿＿＿＿＿＿＿＿や不安の場面を撮った。
    a) 恐怖   b) 強符   c) 恐布   d) 教父
    Completely unsure   1   2   3   4   5   *Completely certain*

18. 円高の影響で企業のこれまでの＿＿＿＿＿＿＿＿を見直した。
    a) 凝石   b) 行跡   c) 業績   d) 業積
    Completely unsure   1   2   3   4   5   *Completely certain*

19. 来年、この会社で重要な＿＿＿＿＿＿＿＿を開催すると決まりました。
    a) 回技   b) 懐疑   c) 会議   d) 会犠
    Completely unsure   1   2   3   4   5   *Completely certain*

20. 姉の結婚後の生活は＿＿＿＿＿＿＿＿になったと聞きました。
    a) 幸幅   b) 幸福   c) 降伏   d) 考服
    Completely unsure   1   2   3   4   5   *Completely certain*

21. その男の子の家まで行くと＿＿＿＿＿＿＿＿の一人が出てきました。
    a) 強大   b) 京台   c) 兄弟   d) 兄第
    Completely unsure   1   2   3   4   5   *Completely certain*

22. 週末から降り続いた大雨で川の水が＿＿＿＿＿＿＿＿した。
    a) 旧蔵   b) 急増   c) 級象   d) 急憎
    Completely unsure   1   2   3   4   5   *Completely certain*

23. 入社の競争率が＿＿＿＿＿＿＿＿の会社でも働きたい気持ちがある。
    a) 戦梅   b) 千培   c) 千倍   d) 専売
    Completely unsure   1   2   3   4   5   *Completely certain*

24. 夜間は病院などの＿＿＿＿＿＿＿＿の場所では静かに話しましょう。
    a) 交響   b) 公供   c) 校強   d) 公共
    Completely unsure   1   2   3   4   5   *Completely certain*