

Doctoral Dissertation

Academic Year 2015

**Phonological Representations
of the Japanese Language:
Levels of Processing and Orthographic Influences**

Graduate School of Media and Governance

Keio University

Miyoko Nakamura

CONTENTS

Acknowledgments	iv
Abstract (English)	vii
Abstract (Japanese)	ix
List of Figures	xi
List of Tables	xiii

Chapter 1: Introduction	1
1.1 Phonological Representations	3
1.2 Orthographic Influences	5
1.2.1 From Sound to Letters: from phonology to orthography	5
1.2.2 Orthographic Influences on Speech Processing	6
1.3 The Objectives of the Present Study	8
1.4 Approach	8
1.4.1 General Approach	8
1.4.2 Specific Approach: Levels of Speech Processing	9
1.5 Dissertation Outline	14

Chapter 2: Related Studies	16
2.1 Japanese Phonology and Writing System	16
2.1.1 Japanese Phonology on Syllable and Mora	16

2.1.2	Japanese Phonology and the Writing System	18
2.2	Research Background of Linguistic Awareness	22
2.2.1	Metalinguistic Ability and/or Linguistic Awareness	22
2.2.2	Language as an Object: Metalinguistic Thinking	24
2.2.3	Phonological Awareness and Phonemic Awareness	26
2.2.4	Previous Studies on Phonological Awareness in Alphabetic Language Speakers	27
2.2.5	Previous Studies on Phonological Awareness in Japanese and Non-alphabetic Language Speakers	28
2.3	Research Background of Speech Segmentation	36
2.3.1	Background of Speech Segmentation and Perceptual Units	36
2.3.2	Previous Studies on Speech Segmentation in Japanese	38
2.3.3	Influence of Writing System on Speech Segmentation in Japanese	40

Chapter 3: Phonological Awareness in Japanese Native

Speakers and the Influence of Orthography 44

3.1	Objectives	44
3.2	Experiment 1: Is the Phoneme Spontaneously Used in a Reversal Task?	45
3.2.1	Method: Participants, Material and Procedure	45
3.2.2	Results and Discussion	48
3.3	Experiment 2: Mora-versus Syllable-based Spontaneous Reversals	50
3.3.1	Method: Participants, Material and procedure	50
3.3.2	Results and Discussion	51
3.4	Experiment 3 : Induced Phonemic Reversal Task	54
3.4.1	Method: Material and Procedure	56
3.4.2	Results and Discussion	56
3.5	General Discussion	61

Chapter 4: Speech Segmentation in Japanese

and the Influence of Orthography 63

4.1	General Framework	63
4.2	Research Questions and Objectives	67

4.3	Experiment 1: Units that Migrate in Japanese: Are They Affected by the Target Script?	70
4.3.1	Method: Participants, Material, and Procedure	71
4.3.2	Results	76
4.3.3	Discussion	80
4.4	Experiment 2: French Listeners	80
4.4.1	Method: Participants, Material and Procedure	80
4.4.2	Results	81
4.4.3	Discussion, Cross-experiment Comparison, and Control Metaphonological Experiment	84
4.5	Experiment 3: Are Migrations in Japanese Listeners Affected by the (Un)availability of Kanji Representations?	86
4.5.1	Method: Participants, Material, and Procedure	87
4.5.2	Results	90
4.5.3	Discussion and Cross-experiment Comparison	93
4.6	General Discussion	94

Chapter 5: Concluding Remarks 101

5.1	Conclusion	101
5.2	Implications for Further Studies	103
5.2.1	Implications from the Research Work in Chapter 3	103
5.2.2	Implications from the Research Work in Chapter 4	107
5.3	Future Directions	109

Bibliography114

List of Publication 140

Notes 141

Appendix 150

Acknowledgments

First and foremost, I would like to express my gratitude to Professor Mutsumi Imai of Keio University for her supervision and support on my dissertation. Without her generosity of succeeding the role of supervisor from Professor Emeritus Shun Ishizaki of Keio University, I could not have accomplished my doctorate. I would like to express my special gratitude to Dr. Ishizaki, Professor Emeritus of Keio University, my ex-supervisor and one of the present advisors, for his continuous encouragement as a cognitive scientist all through my research life in order never to give up my research interests as well as for his support and direction in order to complete my doctorate. Without him, I would not have even challenged for a doctorate. I also would like to express my appreciation to Professor Minoru Shimozaki, and Professor Yuko Nakahama for their precious comments and advice as well as warm encouragement as advisors at Keio University. I would also like to address my gratitude to Professor Yasunori Morishima of International Christian University as an advisor in the beginning of the doctoral course as well as for his help and encouragement all through the way for maintaining my energy and research interests.

The main research works in my dissertation were supported by the Human Frontier Science Program(HFSP) grant (project entitled Processing Consequences of Contrasting Language Phonologies) and had been done in collaboration with the professors and researchers of the Free University of Brussels (Université Libre de Bruxelles) in Belgium with the understanding of the professors at Keio University.

First, my special acknowledgements are due to Professor Régine Kolinsky of Université Libre de Buxelles(ULB) for her rigorous collaboration and precious advice for the main research works, especially in order to submit and make the research to be published in the international journals as well as for helping me from more advanced psychological point of view, including statistical analyses, and improving the research writing into a better-organized and more refined fashion. I extend my acknowledgements to her warm thoughts and compassion even in my private life. My special acknowledgements are also due to Dr. José Morais, Professor Emeritus of the above university(ULB), for giving me a direction and an incentive to this research field as well as his precious comments and advice for the main research works. I would very much like to express my sincere appreciation from the bottom of my heart for their continuous encouragement and support all through the long research period including my private life, especially the physically hard time. Without their further supervision as

specialists and authorities of cognitive speech processing and literacy, I could not have accomplished my challenge for obtaining a doctorate.

My gratitude should also be addressed to Professor Toshio Watanabe, Professor Masaki Suwa, Professor Yoko Hamada, Professor Takaaki Kato, Professor Atsushi Aoyama of Keio University as well as Dr. Michiyo Asano, Dr. Kyoko Teraoka-Sakamoto and all the graduate school project members of Imai's lab. for their precious comments and advice. I would also like to add my gratitude to Dr. Jun Okamoto, Dr. Takehiro Teraoka, and Ms. Nao Sekiguchi-Tatsumi of ex-Ishizaki Lab. for their kind support and comments.

Concerning the research work in Chapter 3, the basis of which is my master's thesis, I would like to express my special appreciation to Professor Emeritus Yuji Suzuki of Keio University, for his supervision and encouragement at that time. I am very grateful to Ms. Carmela Spagnoletti, the third author of the research paper, for her keen insight on linguistic consideration and the overall judgment of the participants' reading level of performance on French language. I thank Dr. Kazuaki Miyagishima and Dr. Kazuhiko Kakehi for their very helpful remarks on an oral presentation of the research. I am also very grateful to Dr. Takashi Muto and Dr. Sumiko Sasanuma, Dr. Katsuo Tamaoka, Dr. Anne Cutler, Dr. Haruo Kubozono, and Dr. Hirofumi Saito, for their precious comments and continuous encouragements, to Frank Rupert for his help in evaluating participants' mastery of the English language, and to Michiko Makita and Naoko Sakurai for their help in judging the Japanese sounds of the stimuli and responses. I also express my special thanks to all the people who volunteered to participate in the experiments carried out in the research.

As for the research work in Chapter 4, I warmly thank Dr. Haruo Kubozono and the entire HFSP team for the fruitful discussions, and particularly, thank Dr. Takashi Otake for providing testing facilities at Dokkyo University. I also thank Dr. Hajime Oiwa and Dr. Shun Ishizaki for providing testing facilities at Keio University. I also thank Ms. Mireille Cluytens and Ms. Fiorella Germeau for materials preparation, Ms. Josiane Lechat for testing help with the French-speaking participants. I would like to express my special thanks to Ava and David Junça de Moraes for analyses of sound files as well as their encouragement in my private life. I express my gratitude to all the participants from the Japanese community in Brussels and from Dokkyo University, Keio University, as well as to volunteers of ULB (Université Libre de Bruxelles) in Belgium (CRCN).

I warmly thank Ms. Laurie Rendon for her highly professional and outstanding outcome of editing and proofreading of my English language of the dissertation.

I would like to extend my gratitude to Professor Emeritus Michiaki Yasumura of Keio University as well as Dr. Kiyoko Uchiyama, Dr. Tomoko Okuma, Dr. Ryuichiro Higashinaka, and Dr. Yayoi Tajima of ex-Ishizaki Lab. for their continuous encouragement and kind advice in the course of obtaining a doctorate. I warmly thank Ms. Michiyo Tsuge for her encouragement and understanding of my situation.

I would also like to express my appreciation to Dr. Kimihiro Nakamura of Kyoto University, Dr. Hiroshi Nakajima of OMRON Corporation and Dr. Satoshi Suzuki of Aoyama Gakuin University to encourage me to maintain my interest for studies of cognitive science as well as Dr. Takayoshi Miyazaki of the University of Tokushima, Dr. Sakae Suzuki of Shonan Institute of Technology, and my neighbors, Drs. Miyeko and George Deux, for their kind advice and compassion who had already experienced the procedure of obtaining their doctorates.

I would like to add my appreciation to the medical doctors, Dr. Kure, Dr. Mizushina, and Dr. Yamakawa and the physical therapists of his clinic as well as Mr. and Mrs. Kizaki and all the staff of oriental physical therapy clinic.

Finally, I would like to express my special thanks to Rev. Yuko Uetake of Iwaki Church, Toshiko Sato, Rieko Uchisugawa, Tatsuya Seki, and all the staff and students of Church School, and Toshiko Toda, Yukio Shimada, late Hiroko Bando, Yoko Sato, and all my Christian brothers and sisters in Fujisawa Church as well as Jean-Pierre, Professor Emeritus van Noppen of ULB at the same time, and Chantal van Noppen, in Champs de Mars Church in Brussels for their spiritual support and comfort. My special thanks also go to Reiko Rimbara, Motoko Kurokawa, Atsuko Asami, Kayoko Onda, Keiko Yoshimura, Susan Denton, Elizabeth Marcott Shaw, and all my friends as well as Tomoko Mizushima, Miyako Tanaka, and all my relatives who had been showing their kind understanding of my challenge for a doctorate. I extend my special thanks to my sister-in-law, Yumi, my sister, Kiyoko, and my niece, Kaai who had encouraged me all through the way until the end, and to my husband, Shiro, for his patience and generosity to understand my challenge.

In the end, I would like to dedicate my dissertation to the memory of my mother and my father as well as my sister, Keiko. My special dedication goes to the memory of my grandfather, Dr. Torao Tanaka.

Abstract (in English)

Phonological representations of the Japanese language were explored in terms of the orthographic influences at the different levels of attention, one at the level of phonological awareness in which listener's attention was paid to speech sounds at maximum and the other at the preattentive level functioning for unattended speech being analyzed only passively. Phonological representations of sublexical units including phonemes, which had been little studied, were explored in the population of Japanese adults who have already acquired both Japanese and alphabetic writing systems. The study employed cognitive psycholinguistic experimental approach.

The first research work examined the nature of the units of phonological awareness. In three experiments, participants were asked to perform a reversal task. The results show that morae are the most prominent units in spontaneous reversal. On the other hand, the participants were perfectly able to manipulate phonemes under request. Yet, detailed analysis of their introspective reports reveals that, most subjects used an interchange of written kana characters instead. The use of such a strategy implies the ability to analyze a kana consisting of a consonant and a vowel into their internal CV constituents. Thus, whereas the nature of the first acquired writing system seems to exert a strong, pervasive influence on the native speaker's metaphonological procedures, such language-specific procedures amount to the ability to perform metalinguistic operations at the phonemic level.

The second research work explored the functional units of speech segmentation in Japanese using dichotic presentation and a detection task requiring no intentional sublexical analysis. Indeed, illusory perception of a target word might result from preattentive migration of phonemes, morae or syllables from one ear to the other. First, Japanese listeners detected targets presented in hiragana and/or kanji. Phoneme migrations did occur, suggesting that orthography-independent sublexical constituents play some role in segmentation. However, syllable and especially mora migrations were more numerous. This pattern of results was not observed in French speakers, suggesting that it reflects native segmentation in Japanese. Further, to control for the intervention of kanji, Japanese listeners were presented with target loanwords that can be written only in katakana. Again, phoneme migrations occurred, while the first mora and syllable led to similar rates of illusory percepts. Overall, these findings suggest that multiple units,

such as morae, syllables, and even phonemes, function independently of orthographic knowledge in Japanese preattentive speech segmentation.

The role of the above basic research works is to contribute to facilitating the acquisition of alphabetic spelling in foreign language learning and teaching for Japanese as well as its application to the computer aided instruction. Another contribution is expected for remediation and support for the dyslexics and persons with reading disabilities.

Keywords:

phonological awareness, preattentive processing, speech segmentation, phoneme, mora, syllable, orthographic influences, Japanese

Abstract in Japanese

要旨

日本語の音韻表象に関し二つの研究により、音声に対して最も注意が払われる音韻アウェアネスのレベル、音声に注意を払わず受動的に音声进行分析する前注意段階の処理レベルで、正書法の影響を明らかにした。従来は研究対象とされなかった音素も含め単語以下の音韻表象を、日本語とアルファベットの双方の書字体系を獲得している成人の日本語母語話者について、認知言語心理学的実証実験により検証した。

研究1では、実験では参加者に反転課題を行うことを求め、音韻アウェアネスレベルでの単位を検証した。結果は自発的な反転を求められると音素や音節と比較してモーラが最も優勢な単位であることを示した。一方で、参加者は要求に応じて音素の反転が完璧にできた。しかしながら、参加者の内省報告によると、大多数の参加者はかな文字の交換という方略を使っていた。これはかな文字の中を子音と母音の構成要素に分析する能力を示唆している。最初に獲得した書字体系の性質が母語話者の音韻アウェアネスの処理に強く影響を及ぼす一方で、そのような言語固有の処理が音素レベルでの言語アウェアネスの操作をする能力を示していることが明らかになった。

研究2では日本語の音声分節の機能的な単位を、両耳異音聴による提示、注意を払わない受動的な聴取による検知課題を用い、音素、モーラ、又は音節が一方の耳から他方の耳へとマイグレーション（移行）が起こった結果、単語を錯聴するという実験で明らかにした。参加者はひらがな又は漢字で提示されたターゲットを検知し、音節、及び特にモーラによる錯聴が最も多かったが、音素レベルでの錯聴も起こり、正書法からは独立した単語以下の構成要素が音声分節の際に役割を果たしていた。このような結果は次の実験でのフランス語母語話者には観察されず、日本語母語話者の音声分節を反映している。更に漢字の心的表象の介入をなくすべく、カタカナでしか表わされない外来語をターゲット語として提示する実験を行ったところ、再度音素レベルでの錯聴が起こり、音節はモーラと同様に錯聴が起こった。これらの結果により、注意を払わず受動的に音声进行分析する前注意段階に於いては、複数の単位、すなわちモーラ、音節、及び音素も、正書法の知識から独立して機能していることを示唆している。

これらの基礎研究の役割としては、日本語母語話者に対する外国語のアルファベット綴りの学習や教授を容易にすること、またその応用としてコンピュータによる学習支援に貢献することである。また、一方で文字の読み書きが困難な難読症の人々の検査や支援に貢献することが期待される。

キーワード：

音韻アウェアネス、前注意段階、音声分節、音素、モーラ、音節、正書法、日本語

List of Figures

Figure 1.1.	Levels of human speech processing.	12
Figure 1.2.	Relation between the levels of human speech processing and two core research works in the dissertation.	15
Figure 2.1.	Standard syllabary (hiragana characters) of modern Japanese, or “gojuonzu” table.	20
Figure 2.2.	Summarized sketch of Japanese phonology and writing system.	21
Figure 3.1.	Experiment 1: percentage of the different types of correct reversal responses.	49
Figure 3.2.	Experiment 2: Percentage of the correct mora or syllable based reversal responses.	53
Figure 3.3.	Percentages of participants’ strategies reported in the post-hoc interview.	57
Figure 3.4.	Example of a participant’s sketch made during the post-hoc interview.	58
Figure 3.5.	Percentages of correct response types and reversal response types of the participants to items including a yoon (V_1CjV_2) as a function of their reported strategy.	59
Figure 3.6.	Percentages of correct response and reversal response types of the participants to items including a friction (e.g., [a f i]) as a function of their reported strategy.	60
Figure 4.1.	Schematic illustration of an experimental trial in the C1 condition for the detection of the target word <i>gendo</i> , presented either in hiragana (i.e., kana, for one group) or in kanji (for the other group).	65
Figure 4.2.	Illustration of the waveforms of the stimuli of a dichotic pair used in the experimental M1 condition of Experiment 1 for the detection of the target word <i>gendo</i> or <i>hairu</i>	75

- Figure 4.3.** Average values of d' scores, separately for each speech constituent and trial type (experimental vs. control), observed on the Japanese-speaking participants examined in Experiment 1. 79
- Figure 4.4.** Average values of d' scores, separately for each speech constituent and trial type (experimental vs. control), observed on the French-speaking participants examined in Experiment 2.83
- Figure 4.5.** Average values of d' scores, separately for each speech constituent and trial type (experimental vs. control), observed on the Japanese-speaking participants examined in Experiment 3. 92

List of Tables

Table 3.1.	Experiment 1: Expected response types (with examples) according to different reference units.	47
Table 3.2.	Experiment 2: Expected response types (with examples) according to different reference units.	51
Table 4.1.	Example of target-absent trials designed for the pair of target words <i>hairu</i> (/hairu/) and <i>gendo</i> (/geNdo/).	66
Table 4.2.	Average percentages of false alarms on target-absent trials and of correct detections on target-present trials, plus average <i>d'</i> scores for each speech constituent tested on the Japanese-speaking participants examined in Experiment 1, separately for the targets presented in hiragana vs. kanji, and overall.	78
Table 4.3.	Average percentages of false alarms on target-absent trials and of correct detections on target-present trials for each speech constituent tested on the French-speaking participants examined in Experiment 2.	82
Table 4.4.	Example of target-absent trials designed for the pair of target words <i>raisu</i> (/raJsu/) and <i>konte</i> (/koNte/).	89
Table 4.5.	Average percentages of false alarms on target-absent trials and of correct detections on target-present trials for each speech constituent tested on the Japanese-speaking participants examined in Experiment 3.	90

Chapter 1: Introduction

“Father McKenzie

Writing the words of a sermon that no one will hear

No one comes near...”

When I hear these lyrics from “Eleanor Rigby,” a well-known Beatles song written by Lennon and McCartney (1966), I am reminded of an article in a Japanese pop-music magazine, in which the writer mentions a boy in her neighborhood who sang the first words as *Hazama Kenji* instead of *Father McKenzie*. Although non-native Japanese speakers may find this strange, for me as a Japanese native speaker it seems very natural and also interesting. Yes, to my ears *Father McKenzie* really does sound like the Japanese male name *Hazama Kenji*. An understanding of the phonological representations of the Japanese language will help to explain why. I will introduce this theme in section 1.1 of this dissertation.

The sound of the English name even reminds me of the *kana* sequence ‘はざまけんじ’(hazama kenji). Why do we remember a letter string when we listen to just speech or spoken words? “*Writing the words of a sermon that no one will hear*”. Yes, when we prepare a speech to be heard by native speakers of our language, we write words in our native writing system. Most of us acquired literacy at an early age and cannot remember how we learned to read and write, or more specifically, how we learned that a speech sound corresponds to a letter. Listening and reading are closely related to each other, as are speaking and writing. The issue of orthographic influences on speech processing will be introduced in section 1.2.

How do we listen to speech? How do humans process speech? Adult native speakers listen to the spoken words without paying much attention to how they listen to their native language speech sounds, and they segment the words with almost no effort when listening to running speech. Is there

any difference between situations when we pay attention to speech and those when we do not? These questions have attracted the attention of cognitive psychologists, especially cognitive psycholinguists and phonologists. The question of how much attention we pay is one of the main focuses of this study, and will be introduced in section 1.4, specifically section 1.4.2.

How do native listeners recognize words from fluent speech, often on the basis of partial acoustic information? Words are recognized through the interaction of sound and knowledge. Thus, human speech processing includes two ways of processing information: a *bottom-up* process and a *top-down* process. A bottom-up process of word recognition starts from acoustic analysis in the auditory organs, and is based on the results of data analysis from the acoustic information present in the speech signal to discover what word is being uttered. It is data driven and passive, whereas a top-down process is concept driven and active, and based on higher-level information. A top-down process uses broad knowledge corresponding to a topic in context; as well as common sense, or concepts for understanding speech stored in the long-term memory; and pragmatic, semantic, and other forms of linguistic information. In other words, at one end of speech processing, there is the pure sound information from what we hear in speech. At the other end, we can use information about a language – for example, by imagining what is a possible word in the language. Listeners are always actively seeking to make sense out of a spoken message. Thus, human word recognition integrates these two processes (e.g., Ishizaki, 2013; Ryalls, 1996).

In spoken word recognition in one's native language, retrieving a mental representation from the speech input is almost automatic and extremely difficult to describe. We need to locate word boundaries in a continuous speech stream and find a single representation corresponding to highly variable acoustic waveforms. Over the last 20 years, speech perception studies have assumed that the acoustic waveform is analyzed into a set of discrete, symbolic sub-lexical units in the representation, which

mediate the mapping between the acoustic signal and the mental lexicon in our mind (Kolinsky, 1998).

I should justify here the reasons why the study of phonological representations of the Japanese language is important and has both theoretical and practical significance.

First, phonological studies on the Japanese language have made a considerable contribution to speech processing studies, because of the importance of a phonological unit, the mora, which is different from the phoneme¹ and the syllable (e.g., Shibatani, 1990) and is an intermediate unit between them. These phonological units will be introduced in the next chapter (see 2.1.1). Indeed, the rhythm of the Japanese language has been described as being based on the mora (e.g., Ladefoged, 1975; Port, Dalby, & O'Dell, 1987; Shibatani, 1990; see more recent proposals in Grabe & Lowe, 2002; Ramus, Nespors, & Mehler, 1999), which functions as a rhythmic unit in the composition of poems like waka and haiku (Shibatani, 1990) and which plays an important role in word formation rules (Itô, 1990; Mester, 1990; Poser, 1990).

Second, the Japanese language is very unique as regards its writing systems. The language has three kinds of traditional writing systems, generally described as being based on three kinds of signs: a logographic system of Chinese origin, namely *kanji*, and two sets of Japanese moraic phonographic systems that stand in a perfect one-to-one correspondence to each other,² namely *hiragana* and *katakana*. In addition, the Roman alphabet, called *romaji*, is used in some cases. The Japanese writing systems will be discussed in more detail in the next chapter (see 2.1.2).

1.1 Phonological Representations

Phonological representations can be articulated from various points of view, which vary in degree of abstractness from surface acoustic-phonetic descriptions to abstract categories. First, phonological representations can

be defined as pure linguistic representations, which are the research interests of linguists, especially phonologists. Such linguistic representations include units ranging from phones or phonemes to intermediate units between a phoneme and a word, such as mora, syllable, and foot. In many languages, even the internal structural representations of a syllable are the concerns of phonologists. Goldsmith (1993) maintained that all theories of phonology can be usefully divided into theories of representations, of levels, and of rules; the theory of representation is the most familiar at present.

Second, phonological representations can be regarded in terms of abstractness³ of the acoustic details of the input, regardless of gender or individual differences. Such phonological representations can be considered mental representations, and these have attracted the interest of many researchers in psycholinguistics.

Vihman et al. (1994) stated that a phonological mental representation is something more abstract than a vocal motor scheme operating in combination with perceptual attention to particular auditory patterns. In child phonology the term “internal representation” is intended to characterize underlying aspects of the child’s understanding and production of speech, and generally is taken to refer to a form of mental “storage” (Vihman et al., 1994, p.17).

However, the child’s internal representation and the linguist’s description are sometimes assumed to be isomorphic. When the linguist equates psychological reality with internal representation, that reality can best be considered roughly equivalent to the psychologist’s term “mental representation,” which is a form of mental processing. Some linguists even have the same view of representations as psychologists have. According to Coleman (1998), to the extent that cognitive representations of linguistic knowledge are “computed” by the human brain, it seems reasonable to demand that the representations proposed in linguistic theories will be at least capable of supporting computation of some kind (e.g., generation, acceptance, transduction, or parsing). Boersma and Hamann (2009) proposed a bidirectional linguistic model with three representations:

If perception is regarded as the mapping from a universal phonetically detailed form to a language-specific phonological structure without lexical access, an inclusion of perception into the grammar model could require three representations. Similarly, many psycholinguists, e.g. McQueen and Cutler (1997), argue that comprehension is not a single module but consists of two sequentially ordered modules for prelexical perception and for word recognition. The listener is confronted with overt phonetic forms from which he/she has to construct an abstract surface structure in a language-specific way.

(Boersma & Hamann 2009, p. 9)

Third, phonological representations should be accounted in terms of a “levels of processing” approach, which will be described in detail in section 1.4.2. In that frame, I studied the nature of phonological representations in Japanese at two levels, and these will be introduced in my dissertation: post-perceptual (phonological) representations will be presented in Chapter 3, and perceptual (phonological) representations in Chapter 4.

1.2 Orthographic Influences

1.2.1 From Sound to Letters: From Phonology to Orthography

One of the remarkable things in the history of human language is the invention of “letters”, especially phonograms (a visual system of language) transcribed from speech sounds (an auditory system of language). Here I focus only on the relation between phonemes in the Japanese language and two kinds of phonogram, namely, kana and alphabets.

Around the 11th century, more than five hundred years before the first alphabet transcription for Japanese, and just after the invention of the Japanese kana writing system, Japanese Buddhist monks developed a kana syllabary called *gojuonzu* based on the Chinese phonology at that time, in order to study the sound and the writing system of Sanskrit, the formal

language of the ancient Indian Buddhist canon. It was not until the 16th century that Japanese works were transcribed using phoneme-based alphabets. However, a phoneme-like description was added to the appendix of the old kana syllabary in 1079 as イ - イ (i.e., /i-i/) ロ - オ (/ro-o/) ハ - ア (/ha-a/), etc. only by kana being used, indicating extraction of vowels by lengthening the syllable. Considering the fact that old Japanese scholars distinguished Mata (vowels) from Taimon (consonants) even though they only used the kana syllabic characters, the notion that Japanese has a single sound as a unit must have been established at least in the domain of Sanskrit study in the Heian Era (11C; Komatsu, 1981). This fact suggests that those ancient scholars who used syllabograms (i.e., syllable-based phonograms) could analyze the internal structure of a syllabogram and extract, at least, a vowel from a syllabic structure.

Japanese was first transcribed into an alphabet (based on Portuguese sounds) by Portuguese missionaries at the end of the 16th century. After the Japanese government closed the door to foreigners other than Dutch people (mid-17C–mid-19C), Dutch traders did a transcription based on Dutch sounds. Then an American missionary did a transcription based on English sounds around the time of the Meiji Restoration (mid-19C) in order to introduce the Japanese language into the Western world (Koizumi, 1978).

However, the role of the phoneme in the speech processing of Japanese native speakers has not received much attention. As we will review in section 2.1, Japanese is not a phoneme-based language, but a syllable or mora-based language. However, is it not necessary to investigate Japanese speakers' awareness of phonemes, and the role of phonemes in speech segmentation?

1.2.2 Orthographic Influences on Speech Processing

Orthographic influences on speech processing have been reported by many psycholinguistics researchers. These influences will be introduced in Chapter 2 (2.2.3 at the level of linguistic awareness, and 2.3.3 at the level of word recognition). Orthographic knowledge, which is related to literacy,

modulates literate people's speech processing in purely auditory tasks. In fact, orthographic representations modulate the pronunciation of novel words in word learning situations. For example, in a study by Rastle, McCormick, Bayliss, and Davis (2011), participants had to learn associations between novel spoken words and pictures, and then learned to spell these words. When asked to name the pictures, they named the novel words more quickly when their spellings were regular rather than irregular (e.g., for initial /k/, k vs. ch; see also Bürki, Spinelli, & Gaskell, 2012). Rastle et al. found significant orthographic effects on speech perception in a situation in which spelling-sound consistency was manipulated, and discussed the results in terms of a highly interactive language system in which there is a rapid and automatic flow of activation in both directions between orthographic and phonological representations (p.1588).

Orthographic representations also help children acquire new oral vocabulary: exposing them to the spellings of new words enhances their memory for both pronunciation and meaning of the words (e.g., Ehri & Wilce, 1979; Ricketts, Bishop, & Nation, 2009; Rosenthal & Ehri, 2008). Spellings are effective because they provide readers with orthographic images that are useful for symbolizing and storing sounds in memory. Spelling-aided sound learning scores were highly correlated with subjects' knowledge of printed words, indicating that this representational process may be used by beginning readers to store printed words in lexical *memory* (Ehri & Wilce, 1979, p.26). Spellings improved students' memory for pronunciations and meanings of new vocabulary words, and enhanced their memory for pronunciation and meanings compared to no spellings. In addition, orthographic knowledge benefited vocabulary learning and diminished dependence on phonological memory (Rosenthal & Ehri, 2008, p.175). Children showed robust learning for novel spelling patterns after incidental exposure to orthography, and stronger learning for nonword-referent pairings trained with orthography. The degree of orthographic facilitation observed in posttests was related to children's reading levels, with more advanced readers showing more benefit from the presence of orthography (Ricketts, Bishop, & Nation, 2009, p.1948).

1.3 The Objectives of the Present Studies

This research aims to explore, from the attention-oriented point of view, how phonological representations of the Japanese language differ according to the levels of speech processing, which will be introduced in section 1.4.

More specifically, the main objective is to examine the relative importance of different phonological units – namely, syllables, morae, and phonemes – at the different levels of processing. One is at the conscious level of speech processing, at which the listener pays the most attention to speech sounds (Chapter 3), and the other is at the early (preattentive) level of speech processing (Chapter 4). Both of these levels will be explained in detail in section 1.4. The two studies presented in this dissertation were conducted in the population of Japanese adults who have already acquired both Japanese and alphabetic writing systems. The research also aims to investigate to what extent these adults are aware of the *phoneme* and its role in speech segmentation.

Another objective is to explore orthographic influences on the phonological representations at those two levels of processing.

The relation between the objectives of the two core research works in Chapter 3 and Chapter 4 and the “levels-of-processing” approach are presented in section 1.4 and summarized in Figure 1.2.

1.4 Approach

1.4.1 General Approach

The present research is situated in cognitive psycholinguistics, which belongs to the broader field of cognitive science, and relates not only to language education and educational engineering as well as computer science, but also to medical science, specifically cognitive neuropsychology and speech pathology.

Psycholinguistics is the psychology of language built on theories that

attempt to explain the mental processes behind speaking, listening, and acquisition by a systematic account of the phenomena. Two approaches are used in the research: experiments and naturalistic observations.

In order to pursue the objectives stated in section 1.3, the present research employed an experimental approach, as did most of the previous studies on phonological awareness and speech segmentation.

1.4.2 Specific Approach: Levels of Speech Processing

In cognitive psycholinguistics, the issue of human speech processing has been studied in two separate research domains: the study of spoken word recognition dealing with the units represented in *perceptual* codes; and that of learning to read and write, which is concerned with *post-perceptual*, intentional, conscious access to grapho-phonological units, or more commonly called “metaphonological ability” (Kolinsky, 1998).

To determine the speech representations used by the speaker-hearer of a language, it is most helpful to use a “levels-of-processing” approach. Morais, Kolinsky, and Nakamura (1996) proposed a three-level model of speech processing that was further developed in Kolinsky’s (1998) stage-processing approach, which includes three processing stages. Each level was assumed to be sensitive to different parameters (see Figure 1.1 for an integration of both models, with modifications that will be described later in this section).

The first and lowest level functions for “perceptual” processing (Morais et al., 1996) as well as for modular, mandatory perceptual processes (Kolinsky, 1998, 2006) including modular operations. This level is influenced only by one’s early linguistic experiences before literacy is acquired (Morais et al., 1996; Kolinsky, 1998). According to the modular theory, this process and low-level representation are not directly accessible to consciousness (e.g., Fodor, 1983; Pylyshyn, 1981).

The second level is for “processes of post-perceptual re-elaboration” (Morais et al., 1996) or “post-perceptual processes” such as recognition and identification (Kolinsky, 1998). Between perception and conscious

recognition or identification, there is space for a level of processing whose function is to re-elaborate the output of the perceptual module – which is itself unconscious – to take account of further relevant knowledge of the same objects or events (Morais et al., 1996) while it contributes to recognition (Kolinsky, 1998).

Both types of processes operate unconsciously, but one might reserve the term “perception” for the modular processes.

The third and highest level is called “conscious access” to the properties represented in the speech percept, in which one is most aware of speech sounds. This level of attention depends on both the format or structure of the representation and the availability of conscious representations (Morais et al., 1996). It is an “intentional analysis process” (Kolinsky, 1998) concerned with explicit analysis of the recognition outputs, and hence entails metaphonological activity or phonological awareness, which is one aspect of linguistic awareness (see Chapter 2).

Both the second and the third levels include the processes that can be influenced by knowledge related to literacy and attentional strategies. In the three-level model, the difference between the two last stages of processing is that the second contributes to recognition, while the third one is concerned only with the explicit analysis of the result of the *output* of the recognition. Therefore, we must suppose that each level of analysis is constrained by the structure of the output of the precedent level, at least in the processes of word recognition. However, as Kolinsky (2006) pointed out,

... there is not necessarily full isomorphism between perceptual processing units and metaphonological units. This issue of isomorphism is related to the question of whether the conceptualization of units (metaphonology) would be only a matter of access; that is, it could be realized only if those units are already represented in the earlier stages of processing (e.g., Fodor, 1972; Morais, Alegria, & Content, 1987; Rozin, 1976). Alternatively, there would be no isomorphism between the way we conceptualize speech and the perceptive processing of it. In fact, the metaphonological representations could be derived from the

perceptive representations by processes that modify the form of the original representations (e.g., Marcel, 1983; Morais, Content, & Alegria, 1987; see above for more detailed discussion). However, the units of perceptual processing would not necessarily correspond to the conscious representations of “metaphonological” units. In these two cases, as well, the use of “direct tests”, implying a discrimination or all the other types of voluntary response based on one unit or the other, is not adequate for identifying the form of perceptive representations. (Kolinsky 2006, pp.35-36, my translation)

Although the model is very rough and incomplete and requires further development, and must be considered as only a working hypothesis, the level-of-processing approach provides a guide to the interpretation of the experimental results in Chapter 3 in terms of conscious access level, and the results in Chapter 4 in terms of perceptual level. Therefore, the same approach of levels of processing was adopted in the present study.

For most people, perception of speech is an unavoidable part of life, and one hears speech even when not paying attention to it. Therefore, I try to reconsider this model from the attention-oriented viewpoint as well. What I propose here in my dissertation is to add one more term, “early stage of preattentive processes,” to be included in the lowest and/or early level of perceptual processes in relation to the study presented in Chapter 4.

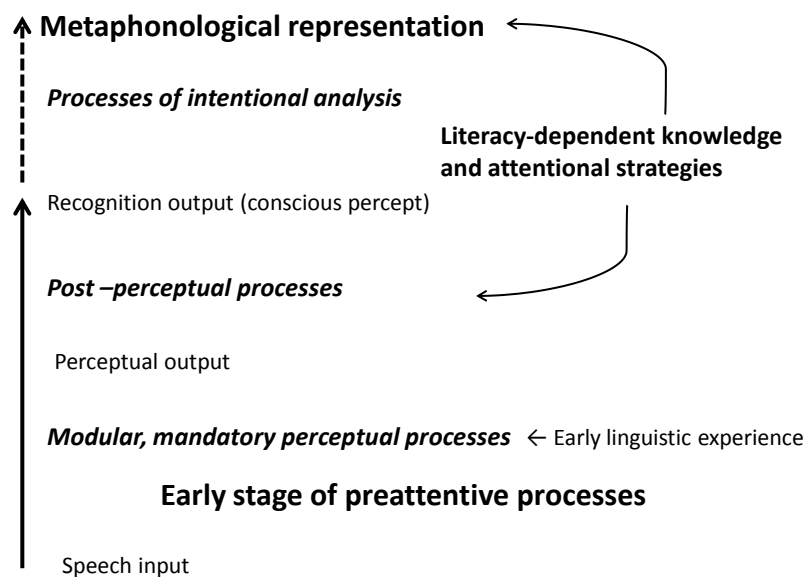


Figure 1.1: Levels of human speech processing.

(Modified from Morais, et al. 1996; Kolinsky, 1998)

First, I should mention the nature of attention. In fact, attention has many senses and “is notoriously difficult to define” (Allport, 2011, p.24). According to Allport, the term “attention” is widely used to denote any and all of the following:

1. At the person level, a relationship between observer and object. At the system level, any such relationship is realized by a corresponding relationship among brain, body, and environment, including an appropriate, *integrated whole-organism brain state*.
2. A *processing constraint* (e.g., “limited capacity”).
3. A control process, or indeed a control *system*.
4. A variety of postulated lower level mechanisms *implementing* such selective control.

(Allport 2011, p.26)

In Broadbent’s (1958) view, attention selects a subset of the information that has been processed by one part of our perceptual system in such a way as to make that information available for processing by a later part of the

system that operates with a smaller processing capacity. The more or less orthodox view, drawing on research by Pylyshyn and colleagues (e.g., Pylyshyn & Storm, 1988) is that attention directly selects a small subset of the objects in the visual scene (Mole, Smithies, & Wu 2011). What can we find in the auditory scene? In his proposed object-oriented episodic record (O-OER) model of working memory, Jones (1993) argued that objects of auditory origin can be formed without conscious control, and separate objects are derived from an auditory stream by relatively simple preattentive processes of streaming and object information.

Returning to the preattentive level of processing (or preattentive processes), in the domain of visual attention, “high-speed serial models” generally assume two stages of processing: a parallel preattentive stage followed by a serial stage of selective attention (Neisser 1967). Early preattentive processes are meant to segregate, group, and otherwise organize the visual array to facilitate the subsequent allocation of attention (Ward, 2001). Treisman (1993) made a further distinction between divided attention and preattention. Preattentive processing cannot directly affect responses or experience; it is an inferred stage of early vision. Before any conscious visual experience is possible, some form of attention is required. Treisman assumed that there was a dichotomy between preattentive and attentive processing levels.

Neisser (1967) proposed the notion of “preattentive control” in visual attention. He maintained that there seem to be two classes of movements that are most often under preattentive control. One of these, which includes head and eye movements, consists of redirections of attention itself. Motion is an effective cue of this sort; when something moves in a portion of the field to which we are not attending, it usually captures our attention almost at once. According to Neisser, “much cognitive activity in daily life is preattentive” (p.92).

With regard to auditory processing, Neisser (1967) distinguished between two levels of processing: a preliminary analysis at a relatively passive preattentive stage provides information, which in turn guides the more active process of synthesis itself. At one level, preliminary

identification of words and other cognitive units is carried out by a passive filter system, which is normally supplemented by an active process of analysis-by-synthesis, suggesting that *this constructive process is itself the mechanism of auditory attention*. Irrelevant, unattended streams of speech are analyzed only by the passive mechanisms, which might be called “preattentive processes” by analogy with the corresponding stage of vision (Neisser 1967, p. 213).

Chapter 4 of my dissertation provides some evidence for this preattentive processing of auditory attention. Therefore I add a “preattentive level” that serves as a sublevel of processing at the first and lowest level (see Figure 1.2). In contrast, the third and highest level of processing, which can be influenced by literacy-dependent knowledge and attentional strategies, is supported by the findings in Chapter 3. Therefore, it is important and significant to view all three levels of processing from the point of view of attention in order that the research works introduced in Chapter 3 and Chapter 4 can deliver consistent attention-oriented accounts.

1.5 Dissertation Outline

The present dissertation consists of five chapters.

In this chapter, I have introduced the general view of phonological representations and orthographic influences, as well as the “level-of-processing” approach and the objectives of the dissertation.

Chapter 2 introduces, first, the Japanese language and writing systems, and second, more specific background of the related studies on phonological awareness as well as speech segmentation.

In Chapter 3, the research experiments in the frame of cognitive psycholinguistics on phonological awareness and the orthographic influences are presented, with method, results, and discussion.

Chapter 4 presents another set of research experiments on speech segmentation and the orthographic influences, also with method, results, and discussion.

Finally, Chapter 5 delivers the conclusion, implications for further studies, and future directions.

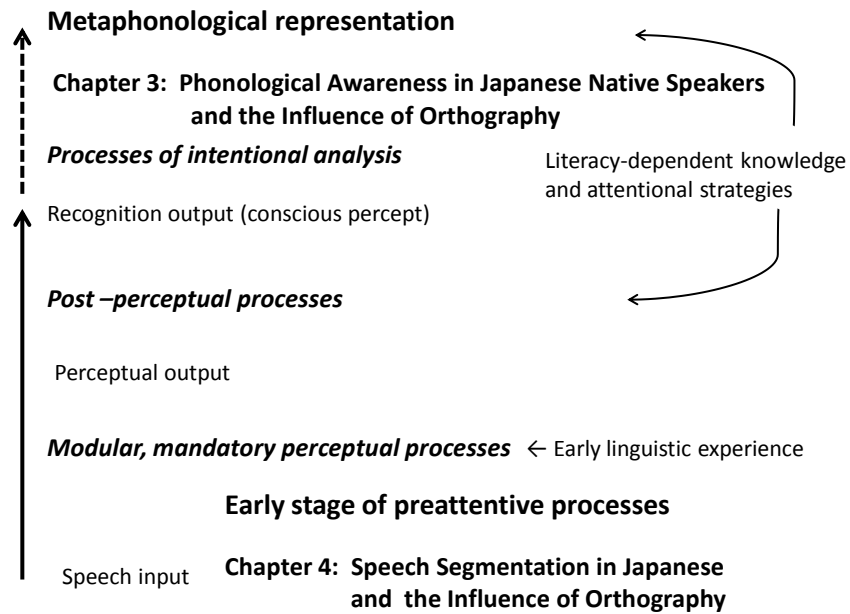


Figure 1.2: Relation between the levels of human speech processing and two core research works in the dissertation.

Chapter 2: Related Studies

2.1 Japanese Phonology and Writing System

2.1.1 Japanese Phonology on Syllable and Mora

As can be seen in Figure 1, modern Japanese language consists of 19 phonemes¹: five vowels (/a, i, u, e, o/), two semivowels (/j, w/), and 12 consonants (/p, b, t, d, k, g, z, s, h, m, n, r/). As can be seen in Figure 1, some phonemes present various allophones (consider for example the different realizations of /h/). In addition, there are some special phonemes, sometimes referred to as archiphonemes, i.e. phonemes for which some features are left unspecified (e.g., Shibatani, 1990), namely /N/ (called *hatsuon*); /Q/ (*sokuon*), and /R/ (*choon*), each of which constituting one mora by itself. The nasal phoneme /N/ is peculiar in that it never appears at the onset of a syllable, and its articulation is homorganic with the following segment. In word final position it is realized as the uvular [N].

As for the syllable structure, Japanese does not admit consonantal clusters at syllable onsets and the sole possible coda consonants are either /N/ or the first member of a heterosyllabic geminate cluster. Besides, phonetic length is distinctive for both consonants and vowels, as illustrated by minimal pairs like /saka/-/saQka/ (*saka*, meaning “slope” - *sakka*, “writer”) and /tori/- /toRri/ (*tori*, “bird” - *toori*, “street”).²

A mora is a noteworthy dimension of standard Japanese phonology. In Japanese, it is a unit that can be represented by one letter of kana and functions as a rhythmic unit in the composition of Japanese poems, e.g. waka and haiku (Shibatani, 1990). Beside its role in the realization of the Japanese accentual patterns, the phonological interpretation of vowel or consonant length is often described by using the notion. It is usually considered to be a subsyllabic unit that can be larger than the phoneme and serves the purpose of measuring syllabic weight (Hyman 1985; Hayes 1989). According to Hyman(1985), “a light syllable has one mora and a heavy

syllable two”, and MaCawly (1977) describes mora as “something of which a heavy syllable consists of two and a light syllable consists of one”.

Although many Japanese words involve simple combinations of a consonant and a vowel (CV hereafter, e.g., *tatemono*, meaning “building”), in which case the number of morae equals the number of syllables (/ta.–te.–mo.–no/)³, the distinction between morae and syllables becomes obvious with other structures. Indeed, there are morae called “special morae” (or “weak” or even “deficient” morae; Labrune, 2012a; b) and also “special phonemes”, given that each of them is constituted by only one phoneme (this is also the case with vowels, but, except for the non-initial *i*, these are considered as “regular” morae [cf. Labrune, 2012a; b] or “autonomous” morae in traditional Japanese phonological theories [e.g., Shibatani, 1990; Tsujimura, 2007]). There are four kinds of special morae: /N/, the final nasal of a syllable, referred to as *hatsuon*; /Q/, the non-nasal moraic consonant referred to as *sokuon*, the first member of a geminate obstruent consonant⁴; /R/, the second member or lengthened part of a long vowel, called *choon*; and the second member of a diphthong (or vowel sequence) in the case of [i], transcribed /J/ in Japanese. These special morae are hierarchically ordered in terms of phonological independence (i.e., whether each of these special morae should be treated as an independent segmentation unit or part of a syllable) as /J/, /R/, /N/, and /Q/, with /J/ and /Q/ displaying the highest and lowest independence, respectively, probably as a function of their sonority (e.g., Kubozono, 1999b; Tanaka, 2000, 2008; Ujihira, 1996)⁵.

In short, the basic syllable structure of Japanese is CV (or V). Japanese phonology does not allow closed syllables, except those in which the coda consonant is either /N/, the syllable final nasal, or /Q/, *sokuon*. Hence, Japanese syllables and morae are mostly isomorphic, but non-isomorphism arises when any of the special morae described above comes just after a CV mora (Kubozono, 1995). The latter part of a syllable (i.e., the special mora) has been interpreted as determining the syllable’s weight. A light syllable consists of one independent mora such as CV or V, while a heavy syllable consists of two morae (an independent CV or V mora plus a special mora,

e.g., Tanaka, 2008). In present study, I will follow other authors (e.g., Kubozono, 1995; 1999a; McCawley, 1968; Shibatani, 1990; Tanaka, 2008) by using the term “syllable” to refer to the heavy syllable. Thus, a word such as *Nippon* (meaning “Japan”) consists of four morae (/ni-Q-po-N/) but to only two syllables (/niQ.poN/).

Both syllables and morae may play a role at the prosodic level, namely in accent assignment in Japanese (e.g., Kubozono, 1998; 1999c; 2003; Shibatani, 1990). For example, Kubozono’s (1999c) careful analysis of the loanword accent rule in Japanese showed that a rule referring only to morae (stating that an accent is placed on the antepenultimate mora, as in e.g. /re.-ba’.-no-N/, “Lebanon”, with ’ denoting the accent) would fail to explain why words like /ro’-N.-do-N/ (“London”) and /so’-Q.-ku.-su/ (“socks”) are accented one mora leftward. In most of the latter cases, the antipenultimate mora is a non-syllabic mora (in the above examples, /N/ and /Q/). Hence, the best and more general rule seems that accent is placed on the syllable containing the antepenultimate mora. Kubozono also illustrates that the syllable plays an indispensable role by referring to various other accentual phenomena in Japanese like compound word accentuation. In addition, he argues that morphological processes also require reference to the syllable structure. For example, in loanword clipping, long words (e.g., *demonstration*, which through the process of vowel epenthesis becomes *demonsutoreesyon* in Japanese) are often shortened to the length of two to four morae (e.g., *demo*). Interestingly, the resultant form must be at least bisyllabic, even when the original loanword begins with a bimoraic syllable (e.g., *pamphlet*, *pa-n.-hu.-re-t.-to*, becomes *pa-n.-hu*). In fact, in Tokyo Japanese, the syllables carry the accent (i.e., the mark of pitch fall), but Japanese accentuation rules must refer to both morae and syllables (Shibatani, 1990): syllables bear accents, but morae bear accentual and phrasal tones.

2.1.2 Japanese Phonology and the Writing System

As regards the writing system, the traditional Japanese system is

generally described as being based on three kinds of signs, namely, logographic system of Chinese origin, kanji, and two sets of Japanese moraic phonographic systems that stand in a perfect one-to-one correspondence to each other, namely hiragana and katakana.⁶ Lexical morphemes such as nouns as well as verbal and adjectival roots are usually transcribed by kanji characters, whereas function words, grammatical particles, and verb or adjective inflections are written by means of hiragana. Katakana is primarily used for loan words and onomatopoeia. In the everyday life, all three kinds of characters are usually combined in the same text. Besides, the Roman alphabet, romaji, is used for the sake of international communication essentially for proper nouns (e.g. train station names, street and highway signs), and more recently for operating the computers and word processors. Romaji transcribes each kana character into an alphabetic sequence.⁷ There are two transliteration system for writing Japanese alphabetically: (1) the *Hepburn system* (*Hebon-shiki* or standard system), a system based on English spelling pronunciation developed by American missionary, James Curtis Hepburn (1815-1911), and (2) the *kunrei system* (*kunrei-shiki* or Cabinet Ordinance system), a phonemic system promulgated by Japanese government (cf. Shibamoto-Smith 1996). The kunrei system is taught at school but it is the Hepburn system which is the most often used in everyday life.⁸

Hiragana as well as katakana are taught at school by using a table in which the characters are arranged in a systematic way. As illustrated in Figure 2.1, columns and rows are organized so that all the kana appearing in the same column begin with the same consonantal sound, except the first column which contains the five vowels, and all the kana of the same row share the same vowel. This reference table is called *gojuonzu* (literally “50-sounds table”) after the original table of old Japanese.⁹ There exists a canonical order of recitation of the table which goes column by column from top to bottom and right to left (/a i u e o/, /ka ki ku ke ko/, etc.). Lists of words and proper names like dictionaries and telephone directories, for instance, respect this order, exactly as English dictionaries and directories are built up according to the alphabetical order.

The Standard Syllabary (hiragana characters) of Modern Japanese, or "Gojuonzu"

CHOKUON										
SEION										
ん	わ	ら	や	ま	は	な	た	さ	か	あ
/N/	/wa/	/ra/	/ja/	/ma/	/ha/	/na/	/ta/	/sa/	/ka/	/a/
		り		み	ひ	に	ち	し	き	い
		/ri/		/mi/	/hi/	/ni/	/ci/	/si/	/ki/	/i/
					[ɸi]		[tʃi]	[ɕi]		
		る	ゆ	む	ふ	ぬ	つ	す	く	う
		/ru/	/ju/	/mu/	/hu/	/nu/	/cu/	/su/	/ku/	/u/
					[ɸu]		[tsu]			[u]
		れ		め	へ	ね	て	せ	け	え
		/re/		/me/	/he/	/ne/	/te/	/se/	/ke/	/e/
っ	を	ろ	よ	も	ほ	の	と	そ	こ	お
/Q/	/wo/	/ro/	/jo/	/mo/	/ho/	/no/	/to/	/so/	/ko/	/o/
	[o]									

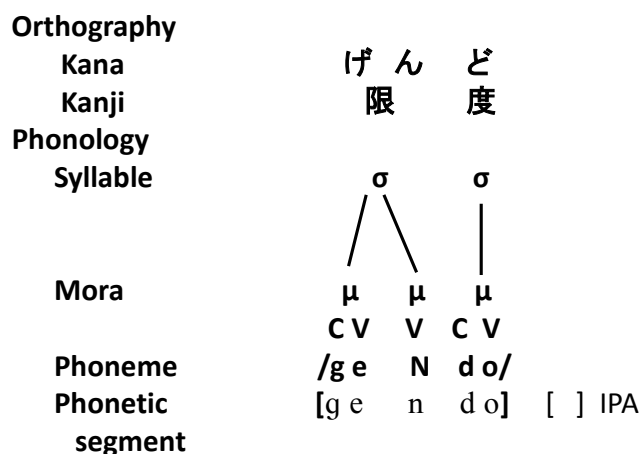
HANDAKUON				
ぱ				
/pa/				
ぴ				
/pi/				
ぷ				
/pu/				
ぺ				
/pe/				
ぽ				
/po/				

DAKUON				
ば	だ	ざ	が	
/ba/	/da/	/za/	/ga/	
		[dza]		
び	ぢ	じ	ぎ	
/bi/	/di/	/zi/	/gi/	
	[dʑi]	[dʑi]		
ぶ	づ	ず	ぐ	
/bu/	/du/	/zu/	/gu/	
	[dzɯ]	[dzɯ]		
べ	で	ぜ	げ	
/be/	/de/	/ze/	/ge/	
		[dze]		
ぼ	ど	ぞ	ご	
/bo/	/do/	/zo/	/go/	
		[dzo]		

YOON											
HAN-DAKUON		DAKUON				CHOKUON					
びゃ	びゃ	ぢゃ	じゃ	ぎゃ		りゃ	みゃ	ひゃ	ちゃ	しゃ	きゃ
/pja/	/bja/	/dja/	/zja/	/gja/		/tja/	/mja/	/hja/	/tja/	/sja/	/kja/
								[ɕja]			
びゅ	びゅ	ぢゅ	じゅ	ぎゅ		りゅ	みゅ	ひゅ	ちゅ	しゅ	きゅ
/pyu/	/byu/	/dyu/	/zyu/	/gyu/		/tyu/	/myu/	/hyu/	/tyu/	/syu/	/kyu/
								[ɕju]			
びょ	びょ	ぢょ	じょ	ぎょ		りょ	みょ	ひょ	ちょ	しょ	きょ
/pyo/	/byo/	/dyo/	/zyo/	/gyo/		/tyo/	/myo/	/hyo/	/tyo/	/syo/	/kyo/
								[ɕjo]			

Figure 2.1. Standard syllabary (hiragana characters) of modern Japanese, or “gojuonzu” table.

Figure 2.1 also shows that the set of 105 kana syllabograms which corresponds to 105 sets of mora of Modern Japanese is divided into two large subsets: simple kana are grouped under the name of *chokuon*, and combinations of two kana are called *yoon*. All *yoon* contain three sounds of which the middle one is the palatal /j/. The first kana is always one of the second row of *chokuon*, namely those ending with the vowel /i/. The second kana, written in reduced size, is always a *chokuon* of the /j/ initial column (for example /ki/ + /ja/ > /kja/ き+や→きゃ). Note that although *yoon* are written by means of two kana, they count as one mora. The summary is sketched in Figure 2.2 with the examples of special morae and kanji/kana transcription.



Notes: Japanese Syllable = CV(V) mora + special mora
(CV (V) + (/J/, /R/, /N/, or /Q/))

Examples of Japanese syllable with different special morae

Phonemic transcription: /ka/ + /J/, /ka/ + /R/, /ka/ + /N/, or /ka/ + /Q/

IPA transcription: [ka i], [ka :], [kan], [kat] ([te])

Kana transcription: かい、 かあ、 かん、 かつ

Kanji (and Kana) transcription: 貝、 お 母 さん、 缶、 勝 (手)

(English meaning) (shell) (mother) (can) ((the first part of) automatic)

Figure 2.2. Summarized sketch of Japanese phonology and writing system.

2.2 Research Background of Linguistic Awareness

2.2.1 Metalinguistic Ability and/or Linguistic Awareness¹⁰

Metalinguistic ability and/or linguistic awareness may be defined as an ability to think about and reflect upon the nature and functions of language, and one of several ‘metacognitive abilities’ (cf. Pratt & Grieve, 1984a,b) as well as an individual’s ability to focus attention on language as an object in and of itself and to evaluate it. “Metacognition refers to an individual’s awareness of strategies and mental activities which carrying out various cognitive processes such as memory, comprehension, learning, and attention.” (van Kleeck, 1987, p.vi, cited from Birdsong, 1989, p.1).

The term “metalinguistics” has been used to describe our conscious knowledge about language, and regulation of its use. Because language use is a cognitive process, a separate, higher-order understanding of language use may thus be termed “metacognitive” (Kurts, 1991). It requires an objective knowledge of the linguistic features and conscious manipulation of them (Bialystok, 1985), and therefore the term “linguistic awareness” is also applied to represent this ability.

Since 1971, a large number of research related to linguistic awareness have been conducted throughout the world. Eight strands of research have been identified by Hutson (1979). A slightly modified version of her topics follows:

- 1) The nature and functions of reading
- 2) The conventions by which we represent language in print
- 3) Definitions of “word,” “sound,” and “sentence,” and differentiation of one term from another
- 4) Understanding of the relationship between words and sentences
- 5) The relationships between words and sounds
- 6) The development of metalinguistic judgment and the ability to hold language as an object of thought and to perform certain logical operations upon it
- 7) The ability to manipulate units of speech intentionally

8) The development of expectations about the literary forms one will encounter in oral and written language.

(cf. Johns, 1984, p.58)

Among the above strands, the research work presented in Chapter 3 is concerned about 7), the ability to manipulate units of speech intentionally.

From the developmental viewpoint, Clark (1978) has reviewed literature on phenomena of language awareness and has classified different types of awareness. The following list contains those different types in their order of emergence and can be interpreted as a preliminary and tentative taxonomy of instances of language awareness:

1. Monitoring one's ongoing utterances
2. Checking the result of an utterance
3. Testing for reality
4. Deliberately trying to learn
5. Predicting the consequences of using inflections, words, phrases, or sentences
6. Reflecting on the product of an utterance (Clark, 1978, p.34)

Hakes (1980) has pointed out important differences between earlier and later forms of metalinguistic performances. While earlier forms seem to arise spontaneously in the ongoing course of conversing, older children are not only able to discriminate spontaneously the properties of utterances but they are also capable of deliberately reflecting on speech. Hakes concludes "that the change in children's metalinguistic abilities is a change in the systematicity and variety of their performances and in the extent to which they can engage in such performances deliberately" (cf. Valtin, 1984, p.208-209).

Further, Levelt, et al. (1978) argues about the causes of linguistic awareness that skills are, by definition, automated activities, in which low-level decisions do not require conscious attention. For any skill, however, there are two obvious circumstances where conscious reflection or decision-making, and therefore awareness, occur. The first one is during *skill acquisition*. The second circumstance where awareness can be observed in the execution of skills is at moments of failure, i.e., when

unexpected or undesired results occur. A similar claim can be found in Piaget (1974) where the child becomes aware in cases of disequilibrium, i.e., where the automatic regulations in performing intentional acts are no longer sufficient to attain a goal (cf. Levelt 1978, p.8).

2.2.2 Language as an Object: Metalinguistic Thinking

As stated in the beginning of the previous section, metalinguistic ability and/or linguistic awareness is defined as an individual's ability to focus attention on *language as an object*. Therefore, it should be also considered from the viewpoint of “decontextualization” or “objectification” of language as people acquire literacy. In fact, according to Birdsong (1989), “Metalinguistic performance in its broadest sense can be understood as any objectification of language” (p.1).

Denny (1991) suggested that the major new thought pattern attributable to literacy is a property called “decontextualization”. Denny found support in Feldman's (1991) work, which showed that Western literacy is not unique in its property of fixing a text and permitting interpretation of it. Feldman pointed out that literacy *facilitates* memory for texts and interpretations, enabling the reservation of more textual material and the development of a longer sequence of commentaries upon commentaries (cf. Denny, 1991, p.77).

Olson (1991) suggested that literacy allowed the “objectification” of language and consequently the development of formal thought skills. He described how writing makes language into an object of such reflection. Writing is used for representing language; it makes it possible to reflect on – to become aware of – language, and thus literacy affects thought. In dealing with written language, one is simultaneously aware of two things: the world *and the language*. Writing takes language for its object, and just as language is a device for “fixing” the world in such a way as to make it an object of reflection, so writing “fixes” language in such a way as to make it an object of reflection. This objectification of language through writing adds to the already existing set of devices for turning speech into an object

of discourse (Olson, 1991, p.266).

The property stated above is similar to the Piagetian term “decentration” used by some authors (e.g., Hakes, 1980; Lundberg, 1978) to refer to the ability to pay attention to (and hence be able to analyze) the expressive or phonological properties of speech while disregarding meaning. The central aspect of linguistic awareness is an *attention shift* from content to form, the ability to make language forms opaque.

Both young children (e.g., Berthoud-Papandropoulou, 1978) and illiterate adults (Kolinsky, Cary, & Morais, 1987) present difficulties in this domain, asserting for instance that (a drawing of) a cat has a longer name than a butterfly – because an actual ‘cat’ is bigger than an actual ‘butterfly’. Usually by the age of six or seven, words are seen as labels that correspond to things, and as such, have an independent existence. When asked to produce long or short words, children of seven and above appear to situate word-length at the level of the graphic representation of the signifier. A general developmental trend in metalinguistic thinking appears in the sense of a reflection on the nature of language itself. Metalinguistic thinking first focuses either on the elements of reality designated by the linguistic units or on the speaker, whether in the characterization of words, in the judgments and production of words, or in the segmentation of sentences. Becoming aware of what goes on when one names objects or action, utters a command, or reads or writes, the child detaches signifiers from their signifieds and starts to reflect on the correspondence between words and things. At this level, words become composite elements that have status proper in their graphic or phonic substance. Simultaneously, the child becomes aware of the possibility of analyzing his own (or others’) utterances into smaller units. At the highest level observed in the present experiments, the children were also able to make explicit the linking of words as signifiers to what they signify, and the links that join words together into groups of units in general discourse (Berthoud-Papandropoulou, 1978, p.58).

2.2.3 Phonological Awareness and Phonemic Awareness

Phonological awareness¹¹ is the very large and heterogeneous set of *conscious representations that are acquired by focusing attention on the speech percept*. This concept must be unpacked by specifying the phonological properties the individual is aware of. (e.g., Morais, 1993). Different forms of phonological awareness are presumably involved; in judging whether two nonsense utterances are identical or different, in deciding which of two evoked name is longer, in appreciating or using rhymes and alliterations (e.g. Bryant, et al., 1990), and in analyzing short utterances into syllables and phonemes (e.g. Byrne and Fielding-Barnsley, 1990).

Phonemic awareness is one of the metaphonological abilities; that is, it entails phonological awareness at the phonemic level as well as the set of conscious representations of the individual phonemes of the language. To date, research on phonemic awareness has focused mostly on alphabetic literacy in alphabetic language speakers (see section 2.2.4).

The focus on problems with phonemic segmentation as a potential cause of reading disabilities can be dated to 1970, when Isabelle Liberman gave an invited presentation to the Orton Dyslexia Society (Mann, 1991). Liberman maintained that reading in the alphabetic system required awareness of the phonemic structure of the language (Liberman, 1971). The same statement was made in the following years by other outstanding researchers such as Mattingly (1972) and Elkonin (1973). The idea, which may seem obvious today, provided the basis for a revolutionary development in the study of reading acquisition and its failures (e.g., Morais, 1987).

Thus phonological awareness is a major concern of those who are concerned with reading difficulties – not only researchers on literacy acquisition (including illiteracy and reading problems) and dyslexia, but also language teachers, children and parents of children with reading difficulties, and dyslexia patients and their speech pathologists. The study of phonological awareness does have an important role to play in helping

those who are in the process of acquiring literacy or have difficulties with it, and those who assist them. This issue is explored in a set of experiments presented in Chapter 3.

2.2.4 Previous Studies on Phonological Awareness in Alphabetic Language Speakers

Most of the studies on phonological awareness in the past 45 years have been undertaken in countries in which alphabetic languages are spoken (for summary, see Bertelson, 1987; Mann, 1991; Morais, 1991), and have addressed literacy and orthography acquisition as well as dyslexia (e.g., Ellis, 1993, Gathercole & Baddeley, 1993). According to Morais (1991), phonological awareness constitutes a bridge between language and literacy.¹²

Phonological awareness is of considerable importance to speakers of alphabetic languages. In particular, alphabetic literacy plays a critical role in the development of phonemic awareness. Among speakers of these languages, only preliterate children (e.g., Liberman, Shankweiler, Fisher, & Carter, 1974) and adults who have never learned an alphabet (fully illiterate adults, e.g., Morais, Cary, Alegria, & Bertelson, 1979; Morais, Bertelson, Cary, & Alegria, 1986) are not aware of phonemes.

In alphabetic literate people, orthographic knowledge modulates performance in purely auditory phonological awareness tasks, because metaphonological representations are closely linked to reading acquisition (e.g., Adams, 1990; Ehri et al., 2001). According to Seidenberg and Tanenhaus (1979), a *congruency effect* of phonology and word spelling knowledge was found in a purely aural rhyme judgment task. The time taken to decide that two spoken words rhyme was shorter when their spellings were similar (e.g., *toast–roast*) than when they were not (e.g., *toast–ghost*); the opposite holding for negative decisions (e.g., faster for *leaf–ref* than for *leaf–deaf*).

Orthographic influences on phoneme deletion, with faster responses for orthographically matched stimulus-response pairs (e.g., *wage–age*) than for

orthographically mismatched pairs (e.g., *worth–earth*), were reported by Tyler and Burnham (2006). With regard to phoneme counting, Ehri and Wilce (1980) found that the number of letters influenced performance; for example, listeners counted an additional phoneme in *pitch* (/pitʃ/) compared to *rich* (/ritʃ/). Castles, Holmes, Neath, and Kinoshita (2003) found better performance for orthographically transparent phonemes, for which there is a direct correspondence with letters (e.g., /d/ in *dentist*), than for orthographically opaque phonemes, for which there is a complex correspondence between the phoneme to be manipulated and the letters representing it (e.g., /n/ in *knuckle*).

Studies of alphabet literate people show that orthographic representations also shaped explicit phonological judgments of the structure of syllables when aurally blending two consonant-vowel-consonant (CVC) monosyllabic words into a new CVC word (Treiman, 1983). Portuguese adults prefer C/VC blends when the word spellings end with a consonant, as in *bar–mel*, /bar mɛl/, but when the word spellings end with a mute e, as in *cure–pele*, /kur pɛl/, they prefer CV/C blends (Ventura, Kolinsky, Brito-Mendes, & Morais, 2001). Even letter names affect metaphonological judgments: in a phoneme counting task, syllables that are letter names (e.g., [ar]) are judged as containing fewer “sounds” than syllables that are not letter names (Treiman & Cassar, 1997).

Furthermore, the influence of orthographic representations extends beyond sub-lexical units: a congruency effect was observed between lexical tones and their orthographic markers at the suprasegmental level in the Thai language (Pattamadilok, Kolinsky, Luksaneeyanawin, & Morais, 2008). Thus, alphabetic language listeners change the way in which they perform phonological awareness tasks, using their spelling knowledge in purely aural situations when they become literate (Kolinsky, in press).

2.2.5 Previous Studies on Phonological Awareness in Japanese and Non-alphabetic Language Speakers

A matter of concern in psycholinguistic research is the nature of the

constituent units in the Japanese speakers' *conscious* representations of Japanese utterances. What are the *metaphonological* units that Japanese speakers develop for their native language? For the Japanese language, the mora has often been proposed as a basic structural unit (e.g., Sugito, 1989). This is partly due to the fact that the number of kana per word nearly always reflects its number of morae. As Fujimura (1966, p.49) explains, "a sense of the mora is traditionally entailed in the Japanese language. The explicit awareness of morae provides the basis for the metrics of Waka and Haiku, two traditional Japanese short poem styles. The kana character is very close to this timing unit as if it had been created according to it."

Other social activities such as word games, in which both adults and children engage, also suggest that the kana, hence the mora, is a prominent metalinguistic unit (e.g., Imada 1990). In the *sakasa-kotoba* game, people play with "reversed words" (e.g., みよこ *Miyoko* こよみ *Koyomi*), which is actually the literal translation of the name of the game. People play with words mainly (but not exclusively) by reversing the order of kana characters, as shown by the reversal of words containing long vowels (e.g., しろう *Shirô*, [ʃiro:] → うろし *Uroshi*). The game has no fixed and established rule neither for the basis of the reversal operation nor for definition of a winner/loser.

On the contrary, the game of *shiritori* has been much more studied because it has more established rules. In fact, it constitutes a good indicator of what unit is actually used in a metalinguistic activity (e.g., Muto, 1987; Katada, 1990; Hatano & Inagaki, 1992). In this game, players take turns giving a word that begins with the end of the precedingly uttered word. The game is over when the word ends with /N/, since no Japanese word can begin with this phoneme (e.g., *neko*, meaning "cat" → *kodomo*, "child" → *moN*, "gate" → game over). Since playing *shiritori* starts in the preliterate period, one can follow the progression of the chosen unit from children's to adults' rules. Muto's (1987) study revealed that *shiritori* players of about four years of age first refer to syllable-based rules. Then after the beginning of kana literacy acquisition, they move towards kana-based (close to mora-based) strategies during their fifth year of age. Muto also reported

some responses based on the last vowel or mora. Unfortunately, the data do not permit to make the difference between the vowels that are part of a CV mora (e.g., *kisha*, “train” → *ashi*, “foot”) and those that constitute a whole mora (e.g., *tokee*, “clock” → *eki*, “station”). Yet, more evidence (Hatano and Inagaki, 1992) has reinforced the idea that young children progressively turn syllable-based into kana-based shiritori rules as their kana reading competence increases. Hatano and Inagaki (1992) concluded that the emergence of the concept of mora as a basic phonological unit of Japanese is subsequent to the acquisition of kana literacy.

Experimental tasks of segmentation corroborate the hypothesis of a progressive emergence of the mora as the basic segmentation unit after a period in which both the syllable and the mora are used. This is the case, for instance, of Inagaki and Hatano’s (1992) study, inspired by Amano’s (1970, 1986, 1988) and Muraishi and Amano’s (1972) pioneering works.

On the other hand, some experimental studies, such as those of Mann (1986) and of Spagnoletti, Morais, Alegria, and Dominicy (1989), were designed to assess the awareness of morae and phonemes in relation to literacy acquisition. Both studies showed that children as young as first-graders excelled at manipulating morae. As regards phonemes, the children’s performance depended on the task. On phoneme counting, Mann found that the task remained a difficult one even by grade six (75% correct responses, 33% being chance level) although Japanese children were clearly above chance level at least from grade four on. Western alphabetized children are, by the end of grade one, nearly at the same level as the Japanese sixth graders (Backman, 1983; Tunmer, Herriman, & Nesdale, 1988). Spagnoletti et al.’s study highlights the strategy used by Japanese first-graders at phoneme counting (namely, tapping). Indeed, CV and VCV items were almost always tapped incorrectly once and twice, respectively, while V and VN items were almost always tapped correctly once and twice, respectively. This strongly suggests that children were tapping by reference to morae or kana.

On phoneme deletion, Mann found poor scores in Japanese first-graders, with an advantage for [k] over [ʃ] deletion (31% vs. 18% correct,

respectively). Spagnoletti et al.'s first-graders performance was overall better than Mann's ones, but showed the same effect of phoneme type (67% vs. 57% correct for [k] and [ʃ], respectively).¹³ This performance pattern is the opposite of the one found in Western children (Content, 1984), while it is consistent with a kana-based strategy consisting in replacing the initial CV kana by the corresponding V kana standing in the same row of the gojuonzu matrix. Indeed, in the kana syllabary, [k] initial characters are spatially closer to vocalic characters than [ʃ]; moreover, the majority of [ʃ] initial morae are spelled by means of digraphs made of the kana for [ʃi] followed either by that of [ja], [ju], [jo].

From the results of phoneme counting or deletion studies, it thus appears that Japanese first-graders are not aware of phonemes. Nevertheless, Mann (1986) suggested that learning a phonographic system, regardless of whether it represents phonemes or not, can even elicit phonemic awareness. Why she suggested so is because fourth-graders who had not yet learned the alphabet even obtained a higher score than first-graders on phoneme counting. Mann further argued that the improvement of performance of her fifth and sixth graders was in part due to the introduction of romaji (the Roman alphabet) in the academic program. The lessons of romaji, eight hours during the second semester of the fourth grade, cannot be compared to the intensive training that pupils receive in western countries. However, the way romaji is taught might trigger the awareness of submoraic/subsyllabic constituents of speech in some way.

Since the kana syllabary is based on a certain kind of phonemic orientation, the instruction provided in the educational context (either in formal setting (school) or informal setting (home)) and also the later awareness of the kana (syllable or mora) manipulation might elicit and bring out their phonemic awareness.

Some properties of the Japanese writing system might actually give hints at the phonemic structure of speech. First, the arrangement of kana into a table in which all the elements that appear in the same row share the vowel might help the discovery of the phonemic structure. Second, some of the morae represented by kana consist of one phoneme, namely the vowels and

the /N/. Third, diacritics are systematically used to distinguish voiced from voiceless initial consonants. Then fourth, some digraphs imply an inframoraic analysis. It has for example been proposed to relate phonemic awareness to the acquisition of yoon reading and spelling.

Therefore, another proposal would be to relate phonemic awareness to the acquisition of yoon reading and spelling. Endo (1991) showed that the better Japanese preschoolers can read and compose yoon syllables, the better they can detect phonemes. However, Morais claims that her evidence is not conclusive since a relatively good performance on some phoneme detection tests can be obtained by illiterate people, who fail at more stringent test of phonemic awareness such as consonant deletion and phoneme reversal . (Morais, 1991; Morais & Kolinsky, 1994, 1995; Morais, Kolinsky, Alegria, & Scliar-Cabral, 1998) ¹⁴

Further, the characteristics of the Japanese phonological system (see 2.1.1) may have profound implications for the nature of the Japanese speakers' phonological units both in speech production and in speech perception. On the production side, reports of speech errors tend to support either the CV syllable (Kamio & Tonoike, 1979) or the mora (Kubozono 1992) as a basic phonological unit of Japanese from linguistic points of view. Yet Terao (1992) reports that in his study only 9% of errors involved the replacement of the whole mora whereas the remaining cases preserved either the vowel or the consonant. Consequently 91% of Terao's corpus of errors can be interpreted as resulting from an interchange of phonemes.

The hypothesis that non-alphabetic phonographic systems may elicit phonemic awareness if they present some useful alphabetic cues is thus not unequivocally supported. As acknowledged by Mann, her own results are ambiguous: the fourth-graders of her study may have been able to discover a useful strategy during the training phase. As a matter of fact, by adding one tap to the number of kana whenever the item includes a mora that is not a vowel or /N/, a correct score of 88% could be reached.⁶ thus, in the absence of further empirical evidence, Mann's (1986) proposal of a late maturation of the phonemic awareness capacity seems too strong. A more conservative position would be that only the learning of a system that explicitly

represents phonemes elicits awareness of these units.

In this theoretical framework, one may wonder whether learning the alphabetic system after having first learned a non-alphabetic system would elicit the development of conscious phoneme representations up to a reasonable level. To our knowledge, up to now only a few studies addressed this question.

Holm and Dodd (1996), focusing mainly on literacy acquisition in a second language, suggested that the non-alphabetic nature of the first-acquired writing system may have long-lasting detrimental effects on both phonological awareness abilities and the capacity to read and write an alphabetically written second language. These authors compared the performance of four groups of Australian students on a series of tasks that assessed phonological awareness and reading and spelling skills in English, a second language mastered by all the participants. One group of participants was coming from Hong Kong, where the alphabetic notation *Hanyu Pinyin* is not taught but where literacy in the Chinese logographs is achieved through the use of a *look and say* approach with the whole character as the basic unit. A second group of participants was coming from the People's Republic of China, where children are first introduced to Pinyin as a transitional alphabet for learning literacy of Chinese logographic system; participants of the third group were coming from Vietnam, where the first written language is an alphabet that uses Roman characters; the fourth group included Australian participants whose first language was English. Holm and Dodd showed that the Hong Kong students had limited phonological awareness compared to the other groups, as well as difficulties in processing English-like nonwords. In particular, the Hong Kong participants were much poorer both for counting the number of phonemes in word stimuli with a one-to-many phoneme to grapheme correspondence (like *whistle*, in which the number of letters fails to correspond to the number of phonemes) and for producing spoonerisms on *letter-digraphs* pairs of stimuli like *soft cheese*, relying far more than the other participants on spelling knowledge (e.g., responding *coft sheese* instead of *choft seese*).

The influence of orthography in metaphonological tasks including phoneme judgments is commonly observed in first-language users, both in children (for example, English-speaking children may count one segment more in the pronunciation of a word like *pitch* than in the pronunciation of *rich*, e.g., Ehri, 1984; Ehri & Wilce, 1980; Perfetti, Beck., Bell, & Hughes, 1987; see also Perin, 1983) and in experienced, mature readers (for example, American adults conceptualize the identity of alveolar flaps as a function of their spelling rather than on a phonetic basis; see Ehri & Wilce, 1986; see also other evidence in Treiman, 1985; Treiman & Cassar, 1997). Yet, over-reliance on the spelling of words seems to be observed mainly in cases where phonemic awareness is defective, as in dyslexics (e.g., Campbell & Butterworth, 1985).

Thus, Holm and Dodd (1996) interpreted the over-reliance on spelling as a symptom of the Hong Kong participants' inability to detect phonemes rather than as the use of an inappropriate strategy. Support for this view comes from the fact that the absence of orthographic information (i.e., of a known spelling), as realized in a non-word phoneme counting task, did not ameliorate the Hong Kong participants' performance (about 20% correct). Since in addition they performed as well as the other groups in word reading and spelling but very poorly in reading and spelling non-words involving either frequent or infrequent graphemes (cf. Treiman, Goswami, & Bruck, 1990), they seemed to use neither grapheme-phoneme conversion nor an analogy process effectively. The authors concluded that when the phonological awareness required for literacy in English had not been developed in the first language, people learning English as a second language transfer their skills formed during first language literacy acquisition to English, being limited, in the case of the Hong Kong students, to poor phonological awareness and to a whole-word, visual reading and spelling strategy.

Yet, one may find good alphabetic readers in adults having acquired an alphabetically written second language. So, if one uses as criterion of selection the participants' nonword reading score in their second, alphabetic language, rather than only the nature (alphabetic *versus* non-alphabetic) of

their first-acquired writing system, one may observe that the non-alphabetic nature of the first-acquired writing system does not prevent developing explicit phonemic representations later on, once an alphabetic system has been taught.

As a matter of fact, studying Chinese residents in the Netherlands who came mostly from Hong Kong, de Gelder, Vroomen, and Bertelson (1993) contrasted a “non-alphabetic” group of participants who could only read Chinese characters with an “alphabetic” group of participants who had since their arrival in the Netherlands acquired some basic Dutch reading ability. The authors observed that on phonemic awareness tests (e.g., deleting the initial consonant from Dutch pseudowords) the “alphabetic” participants obtained far better scores than the “non-alphabetic” ones (e.g., for initial-consonant deletion, on the average 69% vs. 14% correct, respectively). As the findings of Holm and Dodd (1996), this result clearly confirms the main conclusion from studies with illiterate adults that explicit phonemic representations do not develop spontaneously, but only when reading instruction is provided with an alphabetic system (see Adrián, Alegria, & Morais, 1995; Morais, Cary, Alegria, & Bertelson, 1979; Morais, Bertelson, Cary, & Alegria, 1986, for data on illiterate adults; see also Read, Zhang, Nie, & Ding, 1986, for data on Chinese alphabetic vs. non-alphabetic literates). Moreover, de Gelder et al.’s results suggest that learning the alphabetic system after having first learned a non-alphabetic system may elicit the development of conscious phonemic representations, at least when the reader has achieved a reasonable level of alphabetic reading competence.

However, since the de Gelder et al.’s (1993) study was run with Chinese participants, its conclusion may be valid only as far as morphographic or logographic first-acquired writing systems like Chinese hanzi and Japanese kanji are concerned. As already been discussed, a moraic system like the Japanese kana does afford its learners clear infra-syllabic, metaphonological units, namely the morae. One may thus wonder whether Japanese adults who learned the alphabetic system after the Japanese kanji and kana would nevertheless develop explicit phonemic representations, or

whether they would still rely on morae in analyzing speech explicitly. This issue is examined by a set of experiments in Chapter 3.

2.3 Research Background of Speech Segmentation

2.3.1 Background of Speech Segmentation and Perceptual Units

In spoken word recognition, retrieval of unique mental representations from speech input involves the difficult tasks of locating word boundaries in a quasi-continuous stimulus and finding the single representation that corresponds to highly variable acoustic forms. Many researchers have proposed that these segmentation and categorization problems are easier to solve at the sublexical (information below the word level) than the word level: sublexical representations mediate the mapping between the acoustic signal and the mental lexicon.

In particular, the pioneering work of Mehler, Dommergues, Frauenfelder, and Seguí (1981) on French suggested that the syllable plays a key role in this crucial step. Using a *fragment detection task* in which French-speaking participants had to monitor words for a target fragment like *ba* (consonant-vowel, henceforth CV) or *bal* (CVC), these authors observed an interaction between target structure and word structure, with CV targets like *ba* detected more rapidly in words like *balance* (syllabified as *ba.lance*, CV.CVC) than in words like *balcon* (syllabified as *bal.con*, CVC.CV), whereas the opposite result was found for CVC targets like *bal*, which are detected more rapidly in *balcon* than in *balance*. Thus, French listeners' performance in detecting word-initial target fragments was significantly better when the target corresponded exactly to the first syllable of the target-bearing word than when the target constituted more or less than the initial syllable; this suggests a syllable-based segmentation procedure.

Since then, several studies have shown that the basis of speech segmentation differs according to the listener's native language. Several language-specific speech segmentation strategies have been reported, for

example, one that describes segmentation in English according to lexical stress rather than syllables or feet (e.g., Cutler, Mehler, Norris, & Seguí, 1983; Cutler & Norris, 1988).

Findings on speech segmentation lead to the possibility of universality as well as language specificity. With regard to universality, studies of newborns show that they are sensitive to the number of syllables in words, and that the syllable is the universal unit for representing speech (e.g., Mehler & Christophe, 1995); while studies of illiterates indicate that the automatic extraction of phonetic information is unaffected by literacy, and similar patterns of results in literate and illiterate people were obtained (e.g., Morais & Kolinsky, 2001). On the other hand, language specificity on speech segmentation has also been reported. First, infants' early sensitivity to the rhythmic properties of a language lead them to adopt a speech segmentation strategy appropriate to their language (Mazuka, 2009). The emergence of segmentation abilities differs cross-linguistically as a function of the rhythmic class of the language in acquisition (Nazzi, Bertoncini, & Mehler, 1998; Nazzi, Iakimova, Bertoncini, & Frédonie, 2006).

Evidence indicates that native listeners – even if they employ a preferred segmentation strategy – can call upon different processes when these are useful given the nature of the input or the task. For example, the basis of the segmentation processes differs between situations of stimulus presentation in noisy and quiet conditions or when the listeners are put under attentional load (e.g., in lexical decision making with cross-modal fragment priming: Mattys, 2004; Mattys, White, & Melhorn, 2005; in a word identification task: Mattys, Brooke, & Cook, 2009; in artificial language learning: Fernandes, Ventura, & Kolinsky, 2007; Fernandes, Kolinsky, & Ventura, 2010). Even when there is neither physical noise nor attentional load, whether participants resort to a specific segmentation procedure seems to depend on the precise phonetic and phonological nature of the input. For example, whether or not native speakers of French and English resort to a syllabic-based segmentation procedure depends on the nature of the syllabic boundary and/or task demands. In fragment detection

in French, syllables are less relevant when the pivotal consonant is not a liquid (cf. Content, Meunier, Kearns, & Frauenfelder, 2001, who in addition to the liquids /l/ and /r/ already used by Mehler et al., 1981, also used fricatives or stops as pivotal consonants). Furthermore, in English, when the task taps early, prelexical segmentation processes, syllables do play a role in stimuli that are not ambisyllabic (i.e., those that have clear syllable boundaries; Mattys & Melhorn, 2005).

Those data cast doubt on the idea that a unique mechanism or level of structure is always used in speech segmentation. Several mechanisms may coexist. First, these mechanisms are not necessarily mutually exclusive. For example, a syllabic approach to English is not necessarily at odds with the view that lexical access is influenced by the stress-timed properties of the language. As discussed by Mattys and Melhorn (2005, p. 246), “Because syllables are the domain on which stress is applied, perceiving the speech input as a string of syllables is in fact entirely compatible with, and perhaps even implied in, stress-based segmentation” (see related discussions in Kolinsky, Morais & Cluytens, 1995; Murty, Otake, & Cutler, 2007). Second, several phonological constituents might play parallel roles, because some lower-level structures may (at least transitorily) be useful in constraining the specification of other, higher-level structures (Church, 1987; Frazier, 1987).

2.3.2 Previous Studies on Speech Segmentation in Japanese

In a fragment detection task, Otake, Hatano, Cutler, and Mehler (1993) found evidence for a moraic rather than a syllabic effect. Participants were required to monitor Japanese words for the occurrence of either CV (e.g., *ta*) or CVN (e.g., *tan*, /taN/) targets carried by either CVCVCV words like *tanishi*, which include three morae and three syllables (*ta*–*ni*–*shi*) or CVNCV words like *tanshi*, which include three morae (*ta*–*n*–*shi*, with /N/ as the second mora) but only two syllables (*tan*.*shi*). The targets that mismatched the morae boundaries of the carrier words were difficult to detect; particularly, CVN targets (/taN/) were much faster and easier to

detect in CVN.CV words such as *ta-n.-shi* (in which the nasal coda, /N/, served as a separate mora) than in CVCVCV words like *ta.-ni.-shi* (in which the target is only part of a mora); in the latter case, the target was missed more than 60% of the time (see also Otake, Hatano, & Yoneyama, 1996). This pattern was not observed when the same experiment was run with English- or French-speaking participants, leading Otake et al. (1993) to conclude that Japanese speech segmentation occurs at the boundaries between morae.

A similar effect was replicated later in a word spotting task in which Japanese listeners detected Japanese words embedded in nonsense sequences (McQueen, Otake, & Cutler, 2001): words like *uni* (/uni/, meaning “sea urchin”) were detected more rapidly and more easily when they were aligned with the morae boundaries of the carrying sequence (as in *gya-o-u-ni* and *gya-N-u-ni*) than when they were not (as in *gya-bu-ni*); the latter condition led to 99% omissions (see also Sakamoto’s 2005 results on Japanese children). In addition, Cutler and Otake (1994) found that Japanese listeners detected moraic target phonemes more rapidly than they detected targets forming parts of CV morae; thus, for example, /N/ was detected more rapidly in *kanko* (/kaNko/) than in *kanojo*, and /o/ was detected more rapidly in *taoru* than in *kokage* (see also Otake, Yoneyama, Cutler, & van der Lugt, 1996).

However, the exact status of the mora in Japanese remains controversial, as Japanese listeners also seem sensitive to other types of phonological information. As first proposed by McCawley (1968), the syllable may be important as well; that led McCawley to consider Japanese as a “mora-counting syllable language”. The notion that Japanese listeners are sensitive to syllables is supported by analysis of the language’s poetic structures. Although Japanese poetry is largely based on moraic rhythms, according to some authors, both syllables and morae are involved (e.g., Tanaka, 1999, 2008). The analysis of speech production errors also offers evidence for the importance of the syllable in Japanese (Kubozono, 1989).

Such a view would be consistent with the observation that for Japanese listeners, the relevant metric of similarity (which allows listeners to

identify a distorted word in a pun or to reconstruct a word from a nonword in a word reconstruction task) is based on phonemes, not on morae. It would also be consistent with the fact that the point at which a nonword can be rejected in a lexical decision task is predicted better by the phoneme (relative to the mora) boundary at which it diverges from the nearest word (Cutler & Otake, 2002; see also the work on similarity in rap rhyming and Japanese puns by Kawahara, 2007, and Kawahara & Shinohara, 2009a; b).

2.3.3 Influences of Writing System on Speech Segmentation in Japanese

It is noteworthy that several studies argue literacy effects on speech segmentation. First, as for preliterates, Japanese children, for example, shifted their segmentation preference from a mixture of syllable- and mora-based representations to a predominantly mora-based representation as they learned to read kana letters (Otake, Hatano, & Yoneyama, 1996). Whereas, illiterate study discusses that they show an ability to discriminate between the syllables presenting a minimal pair, i.e. the contrastive pair in meaning with only one different phonetic segment (Morais, Castro, Scliar-Cabral, Kolinsky, & Content, 1987).

Dupoux and Mehler (1992) argued that the intervention of a fast orthographic (or associated metaphonological) strategy of target-to-stimulus matching might account for the Japanese fragment detection results reported by Otake et al. (1993). More generally, Kolinsky (1998) argued that tasks such as fragment detection, word spotting, or phoneme detection involve late representations – either metaphonological (i.e., resulting from an intentional and explicit analysis of speech input into sublexical constituents) or orthographic.

Acquisition of metaphonological, explicit representations of speech – particularly at the level of the phoneme – depends on reading acquisition (specifically on learning to read an alphabet): neither preliterate children (e.g., Liberman, Shankweiler, Fisher, & Carter, 1974) nor adults who have not learned an alphabet (either fully illiterate individuals, e.g., Morais, Cary,

Alegria, & Bertelson, 1979; or individuals literate in nonalphabetic systems, e.g., Read, Zhang, Nie, & Ding, 1986) are aware of phonemes. For instance, they cannot tell that there are three phonemes in a word like *car* and are very poor at manipulating phonemes in deletion (e.g., *car-ar*) or inversion (e.g., *car-rack*) tasks and at performing phoneme detection (Morais, Bertelson, Cary, & Alegria, 1986; see discussion in Morais, Kolinsky, Alegria, & Scliar-Cabral, 1998). The metacognitive, explicit representations involved in such tasks differ from perceptual representations. Indeed, the very same illiterate adults who perform poorly on such tasks perform very well (similarly to literate individuals) in perceptual tasks like phoneme discrimination (e.g., Scliar-Cabral, Morais, Nepomuceno, & Kolinsky, 1997; see further evidence in Morais & Kolinsky, 1994; Serniclaes, Ventura, Morais, & Kolinsky, 2005).

Not surprisingly, as metaphonological representations are closely linked to reading acquisition, orthographic knowledge modulates performance in (auditory) metaphonological tasks, like phoneme counting (e.g., Ehri & Wilce, 1979), phoneme deletion or inversion (e.g., Castles, Holmes, Neath, & Kinoshita, 2003), phoneme monitoring (e.g., Frauenfelder, Seguí, & Dijkstra, 1990) and rhyme judgment (e.g., Seidenberg & Tanenhaus, 1979), as was already introduced in the former section (see 2.2.4 (Previous Studies on Phonological Awareness in Alphabetic Language Speakers)).

Furthermore, orthographic knowledge affects not only metaphonological performance but also auditory word recognition. Indeed, responses are slower in auditory lexical decision (e.g., Ziegler & Ferrand, 1998; Ventura, Morais, Pattamadilok, & Kolinsky, 2004) and auditory semantic and gender categorization (e.g., Peereman, Dufour, & Burt, 2009) to words like *deep* – which include rhymes that can be spelled differently in other words of the test language (e.g., as in *heap*) – than to words with rhymes that can be spelled only one way (e.g., *obe* in *probe*). Hence, orthographic and/or associated metaphonological knowledge may also affect performance in the fragment detection task, in which listeners have to detect an embedded sublexical – and even sub-syllabic or sub-moraic – target (e.g., *ba* in *balcon*; *tan* – /taN/ – in *ta.–ni.–shi*) intentionally.

In fact, the notion that the fragment detection task involves literacy-related knowledge has been demonstrated in a Japanese developmental study that examined the influence of knowledge of the kana script on fragment detection (Inagaki, Hatano, & Otake, 2000). Indeed, in Japanese, one mora usually corresponds to one kana character. There are two types of kana, which are the Japanese phonographic symbols: hiragana and katakana (see 2.1.2 Japanese Phonology and the Writing System). Although their usage differs, as was already introduced in 2.1.2, for both, there is an almost systematic one-to-one correspondence of kana to mora. For example, the four morae of the word *Nippon* (/ni-Q-po-N/) correspond to four kana (phonographic) symbols (にっぽん in hiragana or ニッポン in katakana). In addition to the two moraic phonographic systems, the third, logographic system of kanji represents syllables; for example, the two syllables of the word *Nippon* (/niQ.poN/) correspond to two kanji symbols (日本). In addition, it is worth noting that in Japanese, although the different vowels correspond to different kana (each of which corresponds to one mora), no written sign corresponds to an isolated consonant, except for hatsuon (ん, /N/), which always serves as a separate, special mora (see 2.1.1 Japanese Phonology on Syllable and Mora).

As their orthography does not explicitly represent (at least consonantal) phonemes, Japanese children, and even adults, have difficulties in metaphonological tasks that require them to analyze speech intentionally into phonemes (Mann, 1986; Nakamura, Kolinsky, Spagnoletti, & Morais, 1998; Spagnoletti, Morais, Alegria, & Dominicy, 1989), including in phoneme monitoring, as remarked by Otake et al. (1996).

Therefore, orthography and/or writing system also affects speech segmentation in Japanese, at least as measured through fragment detection. Using this task, Inagaki et al. (2000) compared Japanese preschool children who were not yet able to read with older Japanese children who were able to read kana. They found that only the literate children showed the moraic effect first reported by Otake et al. (1993), which they replicated in adults using their own material. The children's results mirrored those observed by the same authors in a metaphonological task, in which children shifted their

segmentation preference from a mixture of syllable- and mora-based representations to a predominantly mora-based representation as they learned to read kana letters (see also Muto, 1987; Otake & Imai, 2001). In addition to showing the relevance of both syllables and morae as segmentation cues in Japanese, these results thus suggest that in the fragment detection task, the acquisition of kana literacy enhances the relative importance of morae compared with syllables.

Chapter 3: Phonological Awareness in Japanese Native Speakers and the Influence of Orthography

As we saw in Chapter 2, the metaphonological as well as the phonological units of Japanese may differ from those of English or French for at least two reasons, namely, the characteristics of both the Japanese phonology and the Japanese writing system.

The phonological interpretation of vowel or consonant length is often described by using the notion of mora. The mora is usually considered to be a subsyllabic unit that can be larger than the phoneme and serves the purpose of measuring syllabic weight (Hyman, 1985; Hayes, 1989). In Japanese, light syllables (i.e., open syllables containing a short vowel) contrast with heavy ones (i.e., open syllables containing a long vowel, and closed syllables). Stated differently, a short vowel (preceded or not by an onset consonant) counts as one mora, a long vowel counts as two, and each coda consonant counts as an extra mora.¹

3.1 Objectives

The goal of the present study was to collect evidence on the relative prominence of different phonological units, namely syllables, morae, and phonemes, in the conscious mental representations of Japanese speech by Japanese adults who know both mora- and phoneme-based writing systems. Using the methodology of the reversal task inspired by Alegria, Pignot, and Morais (1982), the participants were asked to reverse Japanese-like pseudowords after the presentation of a short series of examples. The choice of this methodology was motivated by the fact that reversal operation is

considered to function as the most difficult metaphonological task (Carrillo, 1994; Yopp, 1988; see also Sasanuma, Ito, Patterson, & Ito, 1996) and consequently seems appropriate for the study of adults. In addition, unlike other difficult tasks such as initial phoneme deletion, it cannot be carried out by using non-phonemic strategies (Content, Kolinsky, Morais, & Bertelson, 1986) and since it implies both segmentation and fusion processes as well as some phonological working memory, it relies on elaborated, stable metaphonological entities rather than on perceptual representations

3.2 Experiment 1: Is the Phoneme Spontaneously Used in a Reversal Task?

We know from previous studies that Japanese people can isolate syllabic or moraic units and use them intentionally with great ease. Therefore it was expected that, when involved in a reversal task on Japanese pseudowords, they would spontaneously refer to syllables or morae rather than to phonemes. This prediction was tested in the present experiment. In addition, we wanted to check whether participants used a third strategy, namely, the reversal of graphic, kana characters. Thus, in addition to a simple V_1CV_2 material, some yoon items (V_1CjV_2) were included in the experiment. As shown in Table 3.1, since yoon are written by means of two kana signs but count as one mora, reversing written, kana units would yield different responses than reversing phonological units.

3.2.1. Method

Participants: The Participants were 20 adult native speakers of Japanese, 19 females and one male, aged 19 to 57 years old (mean age: 31.6), belonging to the middle or upper middle class. Among them, 18 were tested in Brussels and two in Japan. Mostly because of their professional occupation, all the participants were well acquainted with at least one alphabetically

written foreign language (mainly English but also French, German, or Spanish). They all had higher education (up to university) and had received formal education on English for 8.5 years, on the average (ranging from six to fifteen years), starting at junior high school (except for three participants who started during elementary school) up to the university. Nine participants learned French in addition to English, for a mean of 2.9 years (range: 1-9 years). Among them, one started during elementary school. Two participants had also learned Spanish (one for six months, the other for one year), and one participant had learned German (for two years). Beside formal education, eleven participants also have had informal, out of school experience with a western language, mainly during stays out of Japan. Four subjects were teachers of English and one was that of Japanese. The participants' alphabetic reading experience was assessed by a test in which they were invited to read aloud a short newspaper article, either in English or in French, according to their own preference. Their reading was recorded for later analysis. The English reading test was evaluated by an English teacher, who was an English native speaker and agreed that all participants were able to read at a reasonable level of fluency. The French reading test was evaluated by two independent French speaking judges. Reading errors were rare (about 2%). It is worth noting that the two participants tested in Japan did not show any difference with the other participants tested in Belgium, either in terms of their foreign language experience or in terms of their performance in the tests.

Table 3.1. Experiment 1: Expected response types (with examples) according to different reference units.

Stimulus type	Reversal type		
	phoneme	mora or syllable	kana
$V_1\text{-}.CV_2$			
[e mo]	[o me]	[mo e]	[mo e]
え も	お め	も え	も え
$V_1\text{-}.CjV_2$			
[a kjɯ]	[ɯ kja]	[kjɯ a]	[jɯ ki a]
あ きゅ	う きゃ	きゅあ	ゆきあ

Hyphens and periods are used to mark moraic and syllabic boundaries, respectively.

The Japanese signs are the kana transcription of the stimuli and the responses.

Material and procedure : Each participant was tested individually in a quiet room. The instructions were given in Japanese by a Japanese female native speaker (the author). The material, recorded by the same speaker, was presented by means of an audio-tape. The required manipulation was a reversal operation. In order to observe the participants' spontaneous behavior, they were merely asked to reverse the items in the way they liked, without specific instruction or feedback correction, then 3 examples and 14 preparatory V_1V_2 trials were presented to them, which were neutral as regards the nature of the reversed unit.

In the experimental phase, 10 three-phonemes disyllabic bimoraic V_1CV_2 sequences were used. As shown in Table 3.1, this structure would give different responses according to the kind of unit (syllable or mora vs. phoneme) taken as reference. Among these items, two included a yoon (V_1CjV_2), which, as also shown in Table 3.1, may reveal kana-based reversals. With a few exceptions, the items and the expected responses were pseudowords. The experimental items were presented in a quasi-random order so as to avoid either sound or kana similarity between adjacent items.

In order to gain more information on the strategies they used, after completion of the experiment the participants were invited to explain how

they performed the task. This post-hoc interview was recorded for later transcription and analysis.

3.2.2. Results and Discussion

As expected, in the preparatory phase the participants did not show any difficulty for reversing vowels: they performed, on the average, at 99.2% correct responses. The instructions to reverse the sequences were thus well understood. Figure 3.1 presents the percentages of each type of response observed on the experimental material, separately for V_1CV_2 and for V_1CjV_2 stimuli.

The vast majority of answers corresponded to syllable- or mora-based inversions. Only one participant provided a phoneme-based inversion: [ɰdʒa] was reversed as [adʒɰ], but at the same time this participant was wondering whether the syllable- or mora-based response [dʒaɰ] would have been more appropriate.

The responses to the items including a yoon did not pattern differently from those to the other items: inversions of written signs were virtually non-existent (only one case was observed). Since yoon are written with two kana characters but constitute one phonological unit (at both moraic and syllabic levels) for the Japanese speaker, one might conclude that the participants did not refer to written units when they have the opportunity to base their response on the mora or the syllable. Yet, during the post-hoc interviews, most participants (65%) reported that they did refer to written signs. Among them all but two reported that they referred to kana constantly (mostly hiragana). Thus, we cannot rule out the possibility that participants processed the graphical compounds (yoon) as one unit, exactly like the other kana.

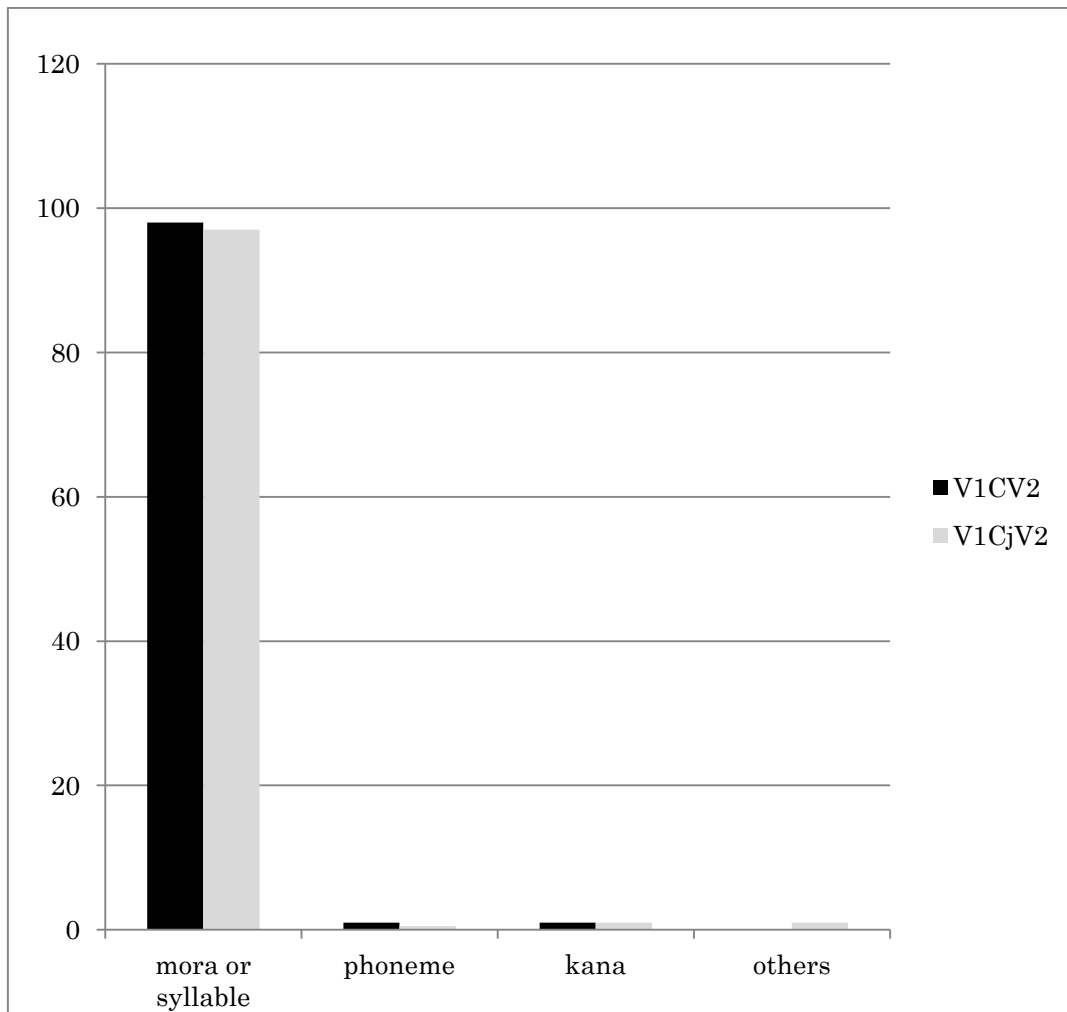


Figure 3.1. Experiment 1: percentage of the different types of correct reversal responses.

In black: V_1CV_2 items, in gray: V_1CjV_2 items.

In any case, it may be concluded that in a task involving an explicit manipulation of speech presented in such a way as to remain unclear on what basis the manipulation must be accomplished, Japanese speakers predominantly refer to the syllabic, moraic, or kana components of their mental representations, but not to phonemes. Phonemes are definitely not the prevalent components of the conscious speech representations of Japanese native speakers.

3.3 Experiment 2: Mora-versus Syllable-based Spontaneous Reversals

In the former spontaneous reversal task, we observed that the vast majority of responses consisted of mora- or syllable-based reversals, as opposed to phoneme-based reversals. While suggesting that the basic metaphonological unit of the special, multiliterate, Japanese population tested here is not the phoneme, the results observed so far have not told us yet which of the syllable or the mora the participants tended to use most spontaneously. The purpose of the present experiment was to specify whether it is the syllable or the mora which is the preferred metaphonological unit of Japanese speakers who, while being native speakers, have a long experience with other, stress- or syllable-based language like English and French. In order to examine this question, used were disyllabic three-moraic sequences, i.e., items in which the syllabic and moraic boundaries do not correspond. Indeed, as shown in Table 3.2, such items would lead to different reversal responses for reversing morae than for reversing either syllables or phonemes.

3.3.1 Method

Participants: The participants were 12 adults, 10 females and two males, aged from 25 up to 58 years (mean: 35yrs), among whom 7 participated in the former experiment. The fresh participants presented the same characteristics of literacy and foreign language experience as the others. Their mean duration of formal education on English (from junior high school till university) was 7 years, ranging from 6 to 9 years; their mean duration of additional formal education on French was 4 years, ranging from 3 to 5 years; 1 participant had learned English at elementary school for 4 years; 3 participants had also informal experience with western languages during stays abroad. Among those who had learned both English and French, 2 had also learned German (1 for 1 year, the other for 2 years), the latter

having also learned Italian for 1 year. Four of them were language teachers (3 English teachers among whom 1 also taught French and 1 teacher of Japanese).

Table 3.2. Experiment 2: Expected response types (with examples) according to different reference units.

Stimulus type	Example	Reversal type		
		phoneme-based	mora-based	syllable-based
V ₁ -N-.CV ₂	[i N pe] いんぺ	[e p N i] え* に	[pe N i] ぺんい	[pe i N] ぺいん
V ₁ -.CV ₂ N	[ʊ se N] うせん	[N e s ʊ] んえす	[N se ʊ] んせう	[se N ʊ] せんう

Hyphens and periods are used to mark moraic and syllabic boundaries, respectively.

The Japanese signs are the kana transcription of the stimuli and the responses when they can be pronounced.

Material and procedure: The test included 20 experimental items, which were constructed by adding to each experimental item used in Experiment 1 the moraic nasal phoneme /N/ either after the first vowel (leading to 10 V₁-N-#CV₂) or at the end of the item (leading to 10 V₁- #CV₂ -N). Before being presented with this material, participants had been presented with V₁-#CV₂ (/oda/ and /azu/) or V₁- #V₂ (/oa/ and /au/), in which syllabic and moraic boundaries do coincide. The other aspects of the material construction and of the procedure were similar to those of the former experiment, except that no post-hoc interview was administered.

3.3.2. Results and Discussion

As can be seen in Figure 3.2, mora-based reversals were far more numerous than syllable-based ones.² Besides, no participant gave any phoneme-based response, despite the fact that at least some of the examples (the V₁- #V₂ ones) could induce phoneme-based reversals. All the

participants gave a majority of mora-based responses. However, two participants out of 12 presented about half of syllable-based responses for the items with the V_1CV_2N structure, and four other participants gave at least one such response. On the contrary, no participant presented a single syllable-based response to V_1NCV_2 items. This response pattern is most probably linked to the fact that mora-based reversals of V_1CV_2N items, although pronounceable without difficulty, would give /N/ initial sequences, which are non-existent in Japanese (see Table 3.2).

The present pattern of results thus shows that Japanese native speakers well acquainted with foreign languages like French and English use the mora to a much greater extent than the syllable when required to perform an explicit manipulation of speech sequences. In addition, while the syllabic structure is available to them, as suggested by the fact that they sometimes resort to it when a mora-based reversal would lead to an illicit sequence (as was the case for the V_1CV_2N items), the majority of participants prefer to produce such illegal sequences than legal syllable-based sequences. These results are thus coherent with those observed on monolingual monoliterates in revealing the mora as the Japanese speakers' most prevalent metaphonological unit.

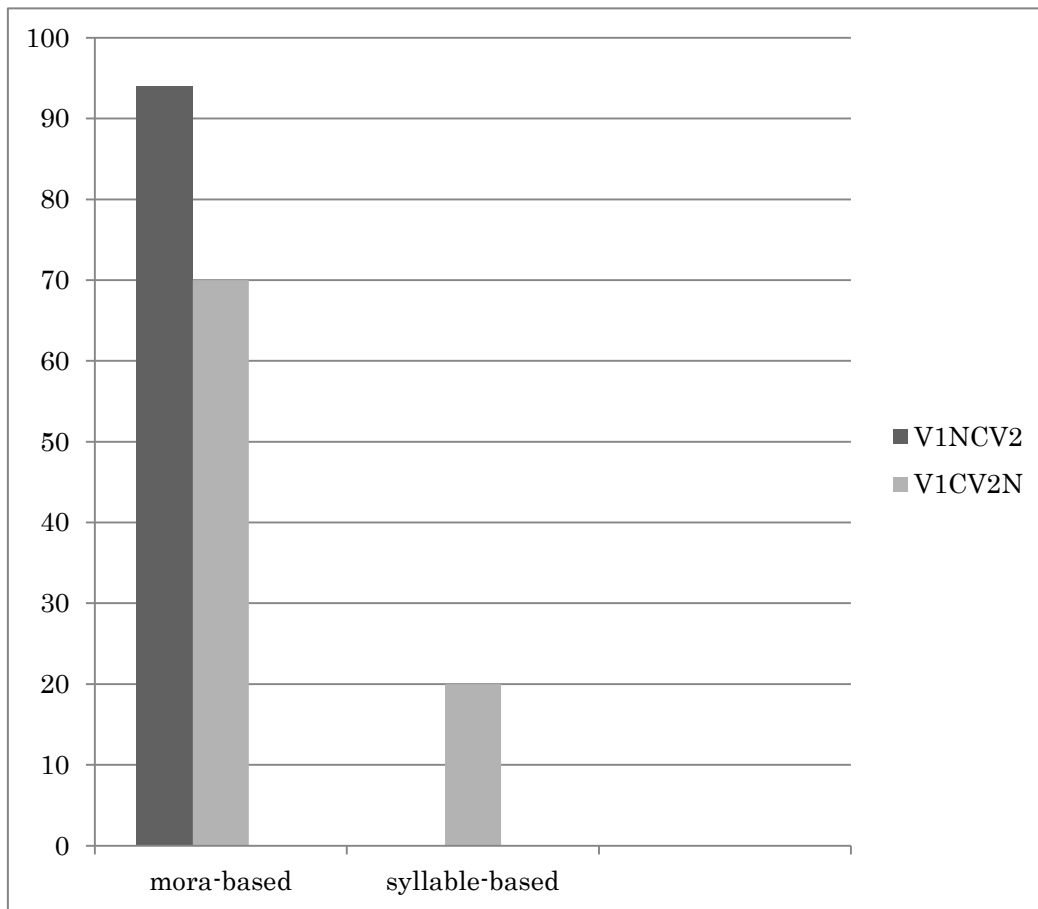


Figure 3.2. Experiment 2: Percentage of the correct mora or syllable based reversal responses. In black: V_1NCV_2 items; in gray: V_1CV_2N items.

3.4 Experiment 3 : Induced Phonemic Reversal Task

The results of Experiment 1 showed that Japanese native speakers, even when knowing an alphabetic, phoneme-based, writing system are not prone to spontaneously refer to phonemes when they are asked to reverse VCV sequences. Coherently, in Experiment 2, while at least some of the examples (the $V_1\text{-}\#V_2$ ones) could induce phoneme-based reversal, a phonemic strategy was never observed. Besides, participants were found to prefer mora- over syllable-based reversals, even when the mora-based operation led to an illicit sequence. Yet, none of these two sets of results implies that Japanese native speakers well acquainted with an alphabetic language would be unable to resort to phonemes when induced explicitly to do so.

Experiment 3 aimed at examining this possibility by inducing participants to perform phoneme reversals. It was also aimed at determining whether the units involved in the reversal operation correspond to phonological or to orthographic units. Indeed, given that the participants were all well acquainted with at least one alphabetic orthography, they were expected to be able to manipulate phonemes under request. But since alphabetic writing is not their first learned writing system, they might tend to resort to written kana or to the alphabetic romaji spelling when performing a phonemic, hence perhaps difficult, task.

A particular interest was taken in checking whether they would use gojuonzu-like table-based strategies, as was already suggested by Spagnoletti et al.'s (1989) study on monolingual Japanese children. A quantitative and qualitative analysis of the participants' responses, combined with examination of their introspective reports, might reveal such procedures.

If the participants were using an orthographic, table-based strategy, the most difficult items to deal with will be those including a yoon, since in this case the table-based localizations and replacement of the units are not straightforward. Furthermore, it was predicted that the participants' response types to the yoon items would differ as a function of their strategies. As a matter of fact, whereas a V_2CjV_1 response (e.g., /apjo/) to a

V₁CjV₂ item (e.g., /opja/) may be obtained on the basis of a visual, table-based strategy (see Figure 3.4), a V₂CiV₁ response (e.g., /aipo/) would be almost impossible to reach using such a strategy.³

The participants' response types to some other special items would also differ according to their strategy. This should be the case with items including a palatal fricative or affricate in the middle position. Indeed, in Japanese a number of consonants are pronounced with a palatal friction when followed by the vowel /i/. For example, before /i/, the [ɕ] sound corresponds to the allophone of the phoneme /s/. What is interesting for the purpose is that in the gojuonzu table (see Figure 2.1), the kana representing [ɕ i] stands in the *i* row of the *sa* column, implying that [ɕ i] belongs to the same phonological category as [sa], namely, the category of all the syllables, or morae, that begin with the same consonantal phoneme. However, the Japanese language allows sequences in which the sound [ʃ] appears before other vowels than /i/, namely before /a/, /u/, and /o/, which is not usually the case with allophonic variants. Japanese scholars agree to describe [ʃi] as the allophonic realization of /si/ whereas the sequences [ʃa], [ʃu], and [ʃo] correspond phonemically to /sja/, /sju/ and /sjo/ (e.g., Shibatani, 1990). Coherently, in kana orthography the sequences [ʃa], [ʃu], and [ʃo] are each transcribed by means of a compound kana, namely, the kana for [ʃi], written in full size, followed by the kana for [ja], [ju], and [jo], respectively in a smaller size than the first one. The same reasoning holds true for the affricates [tʃ] and [dʒ] (see Imada, 1990; Koizumi, 1978; 1989). For example, [tʃi] corresponds to /ti/, whereas [tʃa], [tʃu], and [tʃo] correspond to /tja/, /tju/, and /tjo/, respectively.⁴ Thus, participants using an orthographic, table-based strategy, would tend to show more phonemic-like reversals (e.g., [isa] when presented the item [aʃi]) than participants using the alphabetic romaji spelling. This prediction relies on the fact that the phonemic-like response [isa] may result from a quite simple visual cell interchange in the gojuonzu table (see Figure 2.1), whereas the romaji spelling, which gives an almost allophonic transcription, could lead to the phonetic-like response [ifa].⁵ Similarly, [odʒi] should be reversed as [idzo]

if the reference were table cells, whereas it could be reversed as [idʒo] if the participants were relying on the alphabetic romaji spelling.

3.4.1. Method

Participants. The 20 participants were the same as in Experiment 1.

Material and procedure. Participants were induced to perform a phoneme reversal by presenting them with a series of 10 V_1CV_2 examples which were reversed as V_2CV_1 responses. The 20 experimental trials consisted of 10 V_1CV_2 sequences, 6 V_1CjV_2 yoon items, and 4 items including palatal fricative or affricate in the middle position ([aʃi], [atʃi], [odʒi], and [edʒi]).

The other aspects of the material construction and of the procedure were similar to those of the two former experiments. As in Experiment1, after completion of the experiment, the participants were asked to explain how they performed the task. This post-hoc interview was recorded for later transcription and analysis.

3.4.2. Result and Discussion

When induced to perform a phoneme-based reversal operation, Japanese adults are quite able to do so. Indeed, they obtained 79.5% correct, on the average. Yet, there was a strong effect of item type: performance was clearly superior for the VCV than for the VCjV structure, with on the average 88 vs. 59% correct responses, respectively, $t(19) = 4.47$, $p < .0001$.

The use of orthographic table-based strategies, i.e., of an interchange of table cells instead of a genuine reversal of phonemes, could be responsible for the low scores recorded for the VCjV items. Indeed, for these items the correct answer requires a rather complex cells interchange.

In the post-hoc interview, resort to a strategy referring to the whole or to part of a gojuonzu-like table was acknowledged by the majority (60%) of the participants (see Figure 3.3). Two major table-based techniques that are very much alike were reported. In one technique, the participants recalled either the whole gojuonzu table or the relevant part of it, localized the heard

units, and exchanged the kana cells of the two different columns of the same row. For example, as illustrated by a participant's sketch presented in Figure 3.4, the item /uke/ was analyzed as the /u/ cell of the *a*-column, *u*-row; followed by the /ke/ cell of the *ka*-column, *e*-row. Then the /u/ was substituted by the unit that corresponded to that of the column of the /ke/ but of the same row as the /u/, which gives /ku/. The same reasoning is applied to the /ke/ which is replaced by /e/, the net result being the correct response /eku/. The other technique is similar, except that participants reported that they first prolonged the final morae as to extract the final vowel (e.g., /ukeee/ → /e/). Interestingly, some participants (20%) also mentioned the fact that they sometimes use a table representation that included romaji characters.

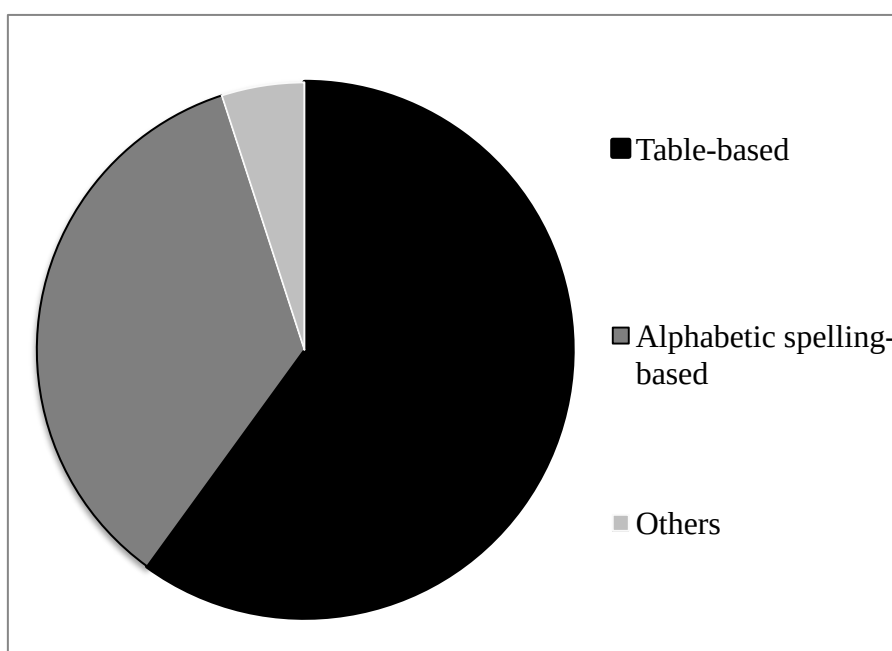


Figure 3.3. Percentages of participants' strategies reported in the post-hoc interview.

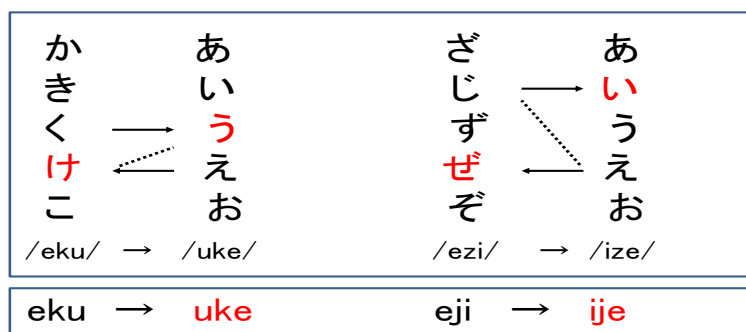


Figure 3.4. Example of a participant's sketch made during the post-hoc interview. (Above: Table(Gojuonzu)-based strategy, Below: Alphabetic (Romaji, Hepburn-system) spelling-based strategy)

The example on the above-left illustrates the kana table-based strategy used to reverse the items /uke/ into /eku/; the example on the above-right shows how /ezi/ was reversed into /ize/ by reference to the kana table (the phonemic transcriptions are added).

Even if a table-based strategy becomes rather complex when involving VCjV items, the mean score observed on the participants having reported to use table-based representations was not lower than that of the 30% of the participants who reported to use exclusively the romaji spelling as visual support (on the average, 57% vs. 51%, respectively). Nevertheless, as predicted, the first, table-based, group only gave V₂CjV₁ responses, whereas the second group, using the romaji spelling strategy, showed 24% V₂iCV₁ and 28% V₂CjV₁ reversals. As may be seen in Figure 3.5, there was a small but significant interaction between the reported strategy and the type of reversal of yoon items, $F(1, 13) = 4.59, p < .05$.

The fact that the medial consonant presented a friction (e.g., [aʃi]) did not reduce significantly the performance level in comparison to the other VCV items (85% vs. 89% correct, respectively, $t(19) = 1.06$). However, as predicted, the response type did vary as a function of the reported strategy:

phonetic-like responses ([iʃa]) were absent in the participants having reported to use a table-based strategy, whereas it was the predominant response for the participants who reported using the romaji spelling. Inversely, phonemic-like response ([isa]), which may result from a quite simple visual cell interchange in a gojuonzu-like table, were the dominant response only in the table-based group of participants. As shown in Figure 3.6, there was a significant interaction between strategy and reversal type for the items including a friction, $F(1, 13) = 89.97, p < .0001$.

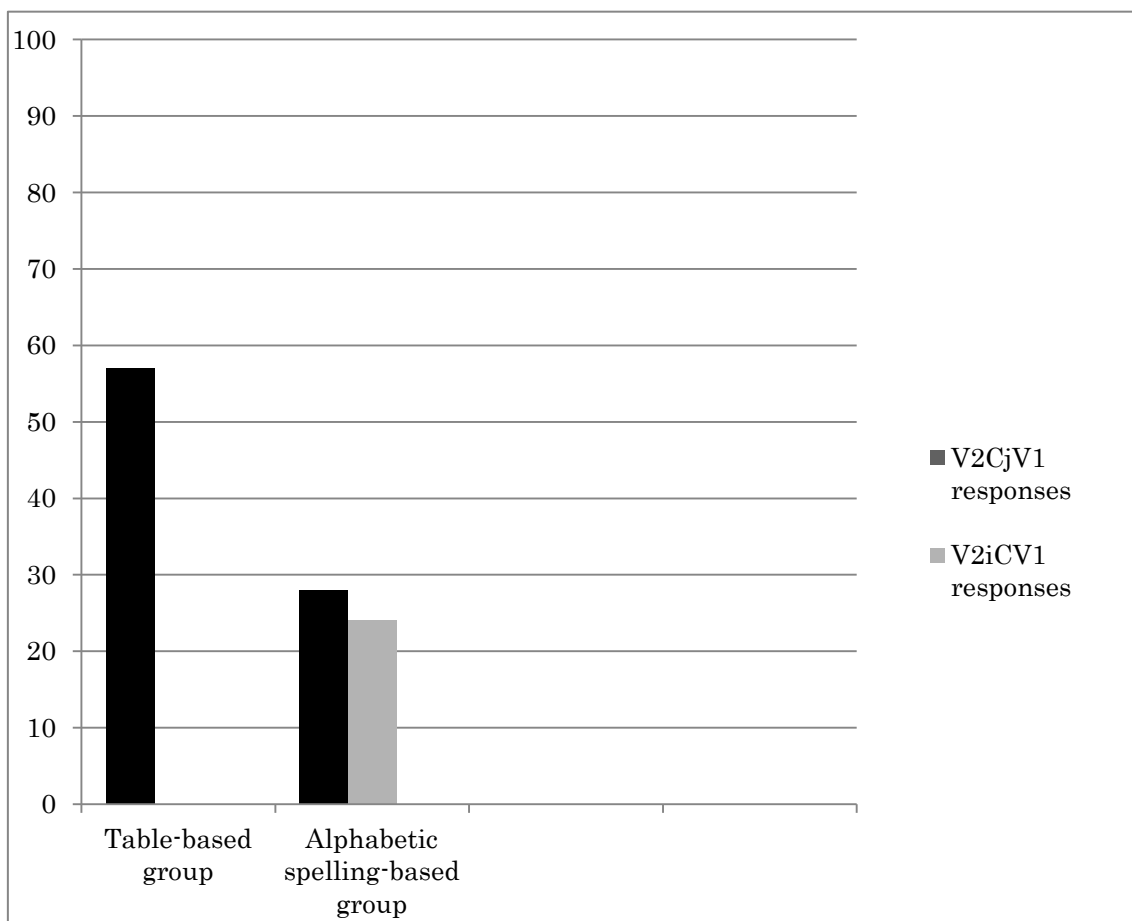


Figure 3.5. Percentages of correct response types and reversal response types of the participants to items including a yoon (V_1CjV_2) as a function of their reported strategy.

In black: V_2CjV_1 responses; in gray: V_2iCV_1 responses.

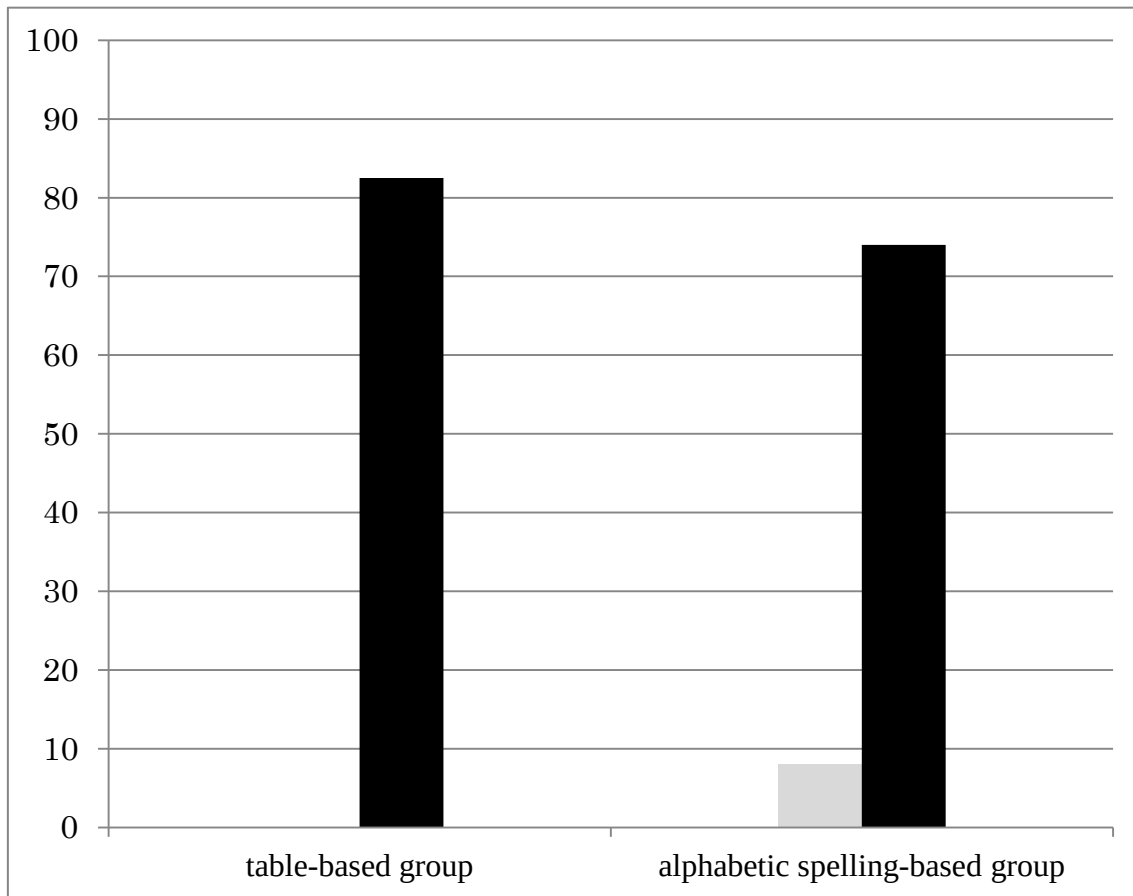


Figure 3.6. Percentages of correct response and reversal response types of the participants to items including a friction (e.g., [aʃi]) as a function of their reported strategy.

In black: phonemic-like responses like [isa]; in gray: a phonetic-like responses like [iʃa].

Besides, it is worth noting that in the group of participants who used the romaji spelling, some responses to these special items revealed a purely letter-oriented strategy. This is for instance the case for the reversal of [atʃi] into [i:ka]. As a matter of fact, given the Romaji transcription *achi*, reversing letters yields *ihca*, with the letter *h* indicating vowel length. Other participants, while responding [ika], explained that they didn't pronounce the "h". Such errors were never observed in the participants who relied on a table-based strategy.

Thus, despite their fairly good ability to perform conscious operations at the phoneme level on Japanese-like utterances when induced to do so by explicit examples, many Japanese adult speakers who are alphabetic literates living in a foreign country prefer a strategy based on the gojuonzu table to a phonemic or an alphabetic strategy. Thus, the nature of the first acquired writing system seems to exert a strong, pervasive influence on the native speaker's metaphonological behavior in their native language.

3.5. General Discussion

The present investigation aimed at examining the nature of metaphonological units used by Japanese adult speakers who also have an advanced knowledge of alphabetic languages like French and/or English. The analysis of reversal performances reveals that the mora is the most prominent unit, but that the syllable can also be used as a spontaneous reference unit when the mora fails to give a satisfactory result.

The participants showed to be also quite able to perform the reversal operation on the phoneme level when explicitly instructed to do so. However, introspective reports revealed that the majority of participants spontaneously imagined a gojuonzu-like table while performing the phoneme-oriented reversal task. As a matter of fact, they mainly reported a strategy which consists of an interchange of kana cells instead of a reversal of phonemes. The use of such a strategy was pointed out in the participants' results by several facts. As predicted, the most difficult items to deal with were those including a yoon, for which no simple cell interchange is possible in a gojuonzu-like table. Furthermore, the participants who reported such a table-based strategy showed qualitatively distinct responses for the yoon items as well as for items including a friction than participants who reported a romaji spelling-strategy. It is worth noting however that in neither case the type of strategy used by the participant affected his/her overall performance level. Even for difficult items like yoons, participants

resorting to a table-based strategy were not poorer at the reversal task than participants resorting to the romaji spelling.

The rather indirect path used by many participants to accomplish the task shows that for them phonemes *per se* are not easily accessible metaphonological units. But at the same time, the identification of the relevant cells to be interchanged in the kana matrix does imply the ability to recognize that a CV kana can be analyzed as a consonant followed by a vowel, and that a vowel can be associated to an initial consonant to give a new CV kana, which amounts to the ability to perform metalinguistic operations at the phonemic level. Thus, despite the fact that the Japanese participants examined in the present work have a rich experience with alphabetically written languages, most of them do not use the same access procedures to the segmental units represented by the French or English orthographies as native learners do, but prefer to use the spelling resources of their native language.

As was introduced in the previous chapter, studies show that newborns are sensitive to the number of syllables in words, and that the syllable is the universal unit for representing speech (Mehler & Christophe, 1995; see also section 2.3.1, p.37). Furthermore, one phonologist claimed the universality of both mora and syllable (Kubozono, 1998) for Japanese speakers/listeners, while other psycholinguists argued for awareness of the subsyllabic unit (mora) and the universality of the mora (Otake & Imai, 2001). In contrast, Morais, Alegria, and Content (1987) maintained that “the ability to deal explicitly with the segmental units of speech is not acquired spontaneously in the course of cognitive growth, but demands some specific training, provided by learning to read and write in the alphabetic system” (p.416). Morais et al. proposed an interactive view of the relationships between segmental analysis and alphabetic literacy. Thus, the reason why the acquisition of phoneme or segmental unit is not spontaneous and needs specific training and/or literacy education has not yet been clarified. This is another matter of interest and it should be examined in a systematic way.

Chapter 4: Speech Segmentation in Japanese and the Influence of Orthography

4.1 General Framework

This chapter presents a study in which the basis of speech segmentation in Japanese was examined. Studies of the Japanese language have made an interesting contribution to research on speech segmentation, because of the importance of a phonological unit, the mora, which is different from (and in a sense intermediate between) the phoneme and the syllable (e.g., Shibatani, 1990). Indeed, the rhythm of the Japanese language has been described as being based on the mora (e.g., Ladefoged, 1975; Port, Dalby, & O'Dell, 1987; Shibatani, 1990; see more recent proposals in Grabe & Lowe, 2002; Ramus, Nespor, & Mehler, 1999), which functions as a rhythmic unit in the composition of poems like waka and haiku (Shibatani, 1990) and plays an important role in word formation rules (Itô, 1990; Mester, 1990; Poser, 1990).

To avoid potential effects of literacy-related knowledge introduced in Chapter 2 (see 2.3.3), some researchers (e.g., Kolinsky et al., 1995; Mattys & Melhorn, 2005) used tasks that do not require listeners to analyze the input intentionally at the sublexical level. They suggested that it is possible to elicit a particular kind of combination error in speech perception, which is called the “migration” of speech constituents (Kolinsky et al., 1995; see also Kolinsky & Morais, 1996). Following the logic of so-called “illusory conjunctions” (Treisman & Schmidt, 1982), we took advantage of the dichotic listening technique to elicit word illusions resulting from the erroneous combination (“migration”) of certain pieces of information presented to one ear with other pieces of information presented to the other

ear (see Figure 4.1 for the schematic illustration of migration experiment).

The migration paradigm originated from studies in the domain of visual perception. Treisman and Schmidt (1982) observed that, when presented with red squares and green circles, participants frequently experienced the presence of a red circle or a green square. According to Treisman and Paterson (1984), such blend errors show that the involved attributes must have been separately registered as independent entities at some (presumably early, preattentive) processing level. This suggestion led to intensive research on blend errors, including those involving written words (e.g., McClelland & Mozer, 1986). By using dichotic stimulus presentation, we adapted this reasoning to the speech domain successfully and observed blending between sublexical portions of the input, namely “migration errors” (Kolinsky & Morais, 1996; Kolinsky et al., 1995). As illustrated in Table 4.1 for the Japanese stimuli used in Experiment 1, two pseudowords were presented simultaneously, one to each ear. The stimuli were constructed so that they could lead to the illusory perception of a target word by “migration” of some sublexical constituent (phonetic features, phonemes, or syllables) from one paired stimulus to the other. Such errors may help to uncover representations of speech constituents, and are the rough perception-domain equivalent of some slip-of-the-tongue errors (namely of spoonerisms, as in saying *The Lord is a shoving leopard* instead of *The Lord is a loving shepherd*) described in production studies (e.g., Dell & Reich, 1980; Fromkin, 1973; MacKay, 1970; Studdert-Kennedy, 1981).

The migration technique has already been used to study speech segmentation in French (e.g., Kolinsky et al., 1995), English (Mattys & Melhorn, 2005; Mattys & Samuel, 1997), and Portuguese (see Morais & Kolinsky, 1994). There is evidence that migration errors occur at a relatively early, preattentive processing level that is unaffected by literacy-associated knowledge. First, phonetic feature blending – the “ancestor” of the migration paradigm – occurs at a level at which the acoustic signal is checked against phonetic categories (e.g., Cutting, 1976; Day, 1968). Second, even illiterate adults, who are not aware of phonemes (e.g., Morais et al., 1979), experience migration of consonantal phonemes

(Morais & Kolinsky, 1994) as well as phonetic feature blending (Morais, Castro, Scliar-Cabral, Kolinsky, & Content, 1987; see also Castro, Vicente, Morais, Kolinsky, & Cluytens, 1995, for data from preliterate children). In addition, the migration technique allows for direct comparison between several speech constituents potentially involved in speech segmentation.

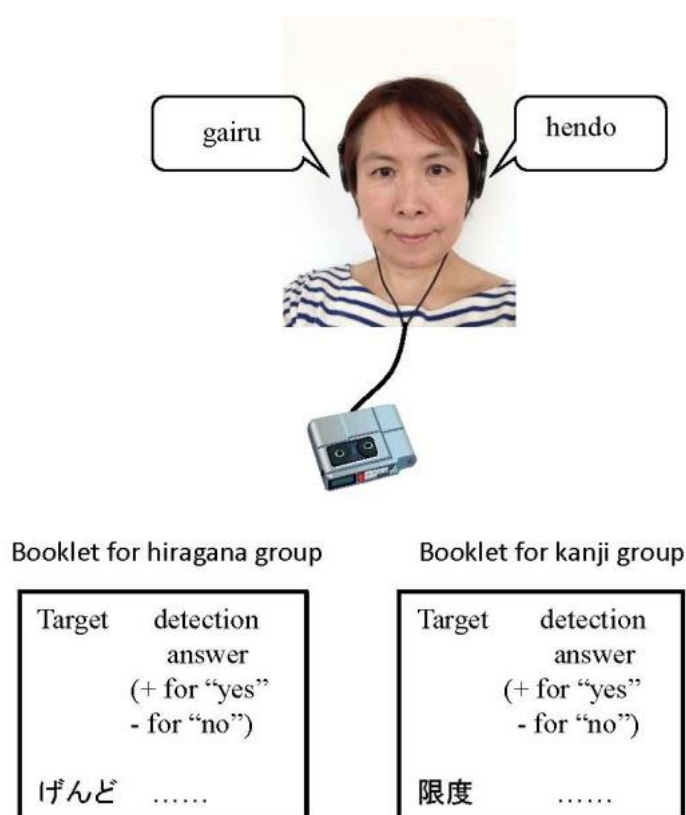


Figure 4.1. Schematic illustration of an experimental trial in the C1 condition for the detection of the target word gendo, presented either in hiragana (i.e., kana, for one group) or in kanji (for the other group).

Table 4.1. Example of target-absent trials designed for the pair of target words *hairu* (/hairu/) and *gendo* (/geNdo/). Letter case signals the distribution (as seen above). Bold characters signal the parts of the experimental trials that have been changed in the control trials. M1('), M2('), and M3(') are the first, second, and third mora, respectively. C: consonant; V: vowel.

Tested constituent	Target	Trial type	Dichotic pair											
			Pseudoword presented to one ear						Pseudoword presented to the other ear					
			M1		M2	M3		M1'		M2'	M3'			
			C1	V1		C2	V3	C1'	V1'		C3'	V2		
C1	hairu or gendo	Experimental	G	a	i	r	u	h	E	N	D	O		
	hairu	Control	G	a	i	r	u	b	E	N	D	O		
	gendo	Control	K	a	i	r	u	h	E	N	D	O		
V1	hairu or gendo	Experimental	h	E	i	r	u	G	a	N	D	O		
	hairu	Control	h	E	i	r	u	G	o	N	D	O		
	gendo	Control	h	O	i	r	u	G	a	N	D	O		
M1	hairu or gendo	Experimental	G	E	i	r	u	h	a	N	D	O		
	hairu	Control (on initial)	G	E	o	t	e	h	a	N	D	O		
	hairu	Control (on final)	G	E	i	r	u	t	o	N	D	O		
	gendo	Control (on initial)	G	E	i	r	u	h	a	Q*	B	A		
	gendo	Control (on final)	B	U	i	r	u	h	a	N	D	O		
Syllable	hairu or gendo	Experimental	G	E	N	r	u	h	a	i	D	O		
	hairu	Control (on initial)	G	E	N	t	e	h	a	i	D	O		
	hairu	Control (on final)	G	E	N	r	u	p	o	e	D	O		
	gendo	Control (on initial)	G	E	N	r	u	h	a	i	B	A		
	gendo	Control (on final)	K	U	Q*	r	u	h	a	i	D	O		

* Q stands for the first member of a geminate consonant

4.2 Research Questions and Objectives

In the present work, exactly the same technique as Kolinsky et al. (1995) was used to explore the functional units in the preattentive segmentation of speech in Japanese. Indeed, by allowing direct comparison between several speech constituents, the migration technique seems appropriate to test the hypothesis that, in addition to morae, both consonantal phonemes and syllables may play roles in Japanese speech segmentation.

Japanese also offers an opportunity to study the role of consonantal phonemes relatively independently of orthography and metaphonology. As no written sign corresponds to an isolated consonant, except for hatsuon (h , /N/), which always serves as a separate, special mora, observing consonantal non-moraic migrations in such a language would constitute strong evidence for a phoneme-based phonological code that is unaffected by orthography and/or metaphonology. Indeed, although for Japanese listeners, morae are more easily accessible to awareness than phonemes (e.g., Mann, 1986; Spagnoletti et al., 1989; Nakamura et al., 1998) and are predominantly used in fragment detection (e.g., Otake et al., 1993), phonemes (like syllables) could nevertheless be involved in early, preattentive speech processing.

To examine such a possibility, compared were the rates of migration errors of the first consonant (henceforth, C1), first vowel (V1), first mora (M1), and syllable in Japanese. As illustrated in Table 4.1 for the target words *hairu* (/ha-J.-ru/, meaning “to go in”, “to come in”) and *gendo* (/ge-N.-do/, meaning “a limit”), manipulated was the distribution of information across pairs of dichotic pseudowords for which the target was not present but could be erroneously blended (henceforth, *target-absent trials*) to compare levels of migration errors involving each of these four types of sublexical constituents. For each constituent, in addition to these *experimental trials*, we also used *control trials* in which the critical information was missing (e.g., for the target *hairu*, the C1 experimental trial was *hendo gairu*, while the C1 control trial was *bendo gairu*; see Table 4.1).

The questions of interest are, first, which migration rates are significant, and second, whether or not migration rates differ between constituents. As a target detection task was used, analyzed was the Signal Detection Theory discrimination index d' (Macmillan & Creelman, 2005). This score is obtained by combining the z-transforms of the false alarm (FA) scores on target-absent trials (i.e., incorrect detections of the target) with those of the *hit* scores (i.e., correct detection scores of the target on dichotic pairs in which the target was actually one member of the pair, henceforth, *target-present* trials); thus, it provides a better, more bias-free indicator of performance than FAs alone.

From the hypothesis that the constituents would be extracted correctly but combined erroneously, it was predicted that the target-absent experimental trials would lead to more FAs than the target-absent control trials and hence be more difficult to distinguish from target-present trials, thus leading to lower d' scores than those thus leading to obtained on the target-absent control trials. Under the hypothesis that phonemes are used in early preattentive speech code independently of orthography, it was predicted that this would hold true for all speech constituents, including C1 and syllables. However, under the hypothesis that segmentation of Japanese predominantly employs both morae and syllables, it was expected to observe an interaction between trial type (experimental vs. control) and speech constituent: a larger effect of trial type was expected for M1 and Syllable compared with C1 and V1. In addition, if the most preferred segmentation procedure were a moraic one, we would observe also a larger effect of trial type for M1 compared with Syllable.

In Experiment 1, the target words were chosen such that there were systematic (except for one case), one-to-one correspondences between the moraic structure and the kana written transcription of the target word and between the syllabic structure and the kanji written transcription of the target word (see illustration in Figure 4.1). For example, the target words *hairu* and *gendo* are written in hiragana with three kana symbols each (はいる and げんど, which correspond to the three morae /ha-J-ru/ and /ge-N-do/, respectively), while in kanji, the same words are each written

with only two characters, 入る (a kanji character plus a hiragana suffix, the so-called *okurigana*) and 限度 (two kanji characters), which correspond to the two syllables /haJ.ru/ and /geN.do/, respectively. Such words were chosen to control for the possible influence of the written transcription. Indeed, the kana transcription may push listeners to rely on the corresponding moraic representations, while the kanji (or kanji plus *okurigana*) transcription may push them to rely on the corresponding syllabic representations.

In Experiment 2, the same material as in Experiment 1 was presented (with oral targets) to non-native (French-speaking) listeners. This allowed us to check whether the results observed in Japanese speakers in Experiment 1 reflect native language segmentation processes rather than material-dependent strategies.

Finally, in Experiment 3, examined was the possibility that the response pattern observed in Japanese participants in Experiment 1 might have been affected by activation of specific mental representations of the words' spellings (i.e., their kanji representations). Although all target words used in Experiment 1 can be transcribed into kana characters only, Japanese people would usually transcribe them by means of either two kanji characters or a single kanji character plus a hiragana suffix. Thus, to detect the auditory presence of such words, Japanese participants might revert to their mental representations in kanji, whatever the actual transcription of the target word.

Hence, the fact that one kanji character often corresponds to one syllable may have favored the syllable as the basis of segmentation in Experiment 1. To check for this possibility, in Experiment 3, used were only loanwords for which there exist no kanji representations (but only katakana ones) as targets.

In this latter experiment, the material was chosen in order also to examine the involvement of the second mora in migration errors. Specifically, examined was the role of the “special morae” /N/ (*hatsuon*) and /J/, [i], the second member of diphthong. These constitute the latter parts of heavy syllables and probably do not have the same phonological

independence as “regular,” “autonomous” morae (e.g., Kubozono, 1999b; Tanaka, 2000, 2008; Ujihira, 1996).

In sum, two main issues were addressed in this investigation: (1) whether in a language such as Japanese – which presents a complex hierarchy of phonological constituents – several of these constituents are functionally relevant at an early, preattentive, stage of processing, and (2) whether orthography and/or metaphonology pervades this processing level, or instead, whether this level is immune to the influence of literacy-related knowledge.

4.3 Experiment 1: Units that Migrate in Japanese: Are They Affected by the Target Script?

In the present experiment, native speakers of Japanese were presented with dichotic pairs and had to detect target words that were presented to them visually, either in kana only (more precisely, in hiragana) for one group of participants, or in kanji (or kanji with a hiragana suffix) for another group (see Figure 10). The migrations of C1, V1, M1, and Syllable were examined; it was predicted that migrations would occur for several speech constituents, including C1, even though single consonants had no isolated orthographic counterparts in the written transcription of the target words (except for /N/, but this cannot be a C1 in Japanese).

The influence of orthography was further examined by comparing the relative rates of M1 and Syllable migrations, checking whether they depend on target script. If orthography affected the segmentation units – as there were almost systematic one-to-one correspondences between the target word’s moraic structure and kana written transcription and between its syllabic structure and kanji written transcription – we would observe more M1 and Syllable migrations in the groups presented with the targets written in hiragana and kanji, respectively.

4.3.1 Method

Participants: All participants were native speakers of Japanese, 61% of them being from the Tokyo area; 28 were tested at Dokkyo University (ages 18–21 yrs; average: 20 yrs; 14 women) and 28 at Université Libre de Bruxelles, Belgium (ages 19–47 yrs, average: 31 yrs; 18 women). The latter were all well acquainted with at least one foreign language (mainly English, but also French, German, or Dutch); their average duration of stay in Brussels was 2.2 years. All participants were paid for their participation. Within each group, half of the participants were randomly assigned to the targets written in hiragana (kana system), and the others received targets written in kanji.

Material: All the stimuli were pseudowords, and all the targets were Japanese words. These words were all intended to include three morae and two syllables, as they were CVVCV, CVNCV or CVQCV sequences.

In addition, a metaphonological test checked whether the target words are indeed considered as including three morae and two syllables by native speakers of Japanese. The test was necessary because the CVVCV sequences included not only the C+*ai* (/aJ/) sequence, constituting a syllable in Japanese, but also the C+*oe* sequence although /oe/ is not considered as a syllable in Japanese phonology (e.g., Fujimura, 2007; Kasuya & Sato, 1990; Kubozono & Ôta, 1998; Saito, 2003; Shirota, 1993). Nevertheless, some researchers admit that it is not clear whether two adjacent non-identical vowels belong to a single syllable or two separate syllables (e.g., Vance, 1987; 2008). In other words, in the first case, a sequence like /oe/ is analyzed (as other VV sequences) as a diphthong.

In the metaphonological test, after having presented some of the spoken target words (*hairu*, *maido*, *soegi*, *moeru*), nine independent naïve native speakers of Japanese (all female, aged 20–49 yrs) were asked to divide the sound, wherever and into how many fragments they thought appropriate. So did we for another six similar CVVCV words (total: 10 words). We observed more than 75% “two” responses, according to which the word seems to

correspond to two syllables; the other responses were 23% “one” responses (i.e., no segmentation) and 2% “three” responses (probably corresponding to mora- or kana-based responses). Thus, they agreed with the idea that the stimuli had a disyllabic structure although these results are somewhat surprising as they reveal a preference for a syllabic mode of explicit, intentional segmentation. This is perhaps because they were induced by the vague instructions given to the participants. This holds true even for the sequences involving /oe/, probably because this is analyzed (as other VV sequences) as a diphthong. Yet, another possibility is that the sequence C+/oe/ is considered as a bimoraic foot, which is also a well-known rhythmic unit playing an important role in Japanese accentuation as shall be discussed later on (e.g., Kondo, 2006; Kondo & Arai, 1998; Kubozono, 1999c; 2006; Mester, 1990; Nasu, 2006; Poser, 1990; Shinohara, 2000; Tanaka, 1998; Tsujimura, 2007).

Sets of pseudoword pairs were chosen so that all experimental trials corresponding to one set would share the same resulting word illusion pair. Each of the pairs of stimuli differed in the particular constituent, for which erroneous perception of the target word would occur because of possible migration from one stimulus representation to the other (see, e.g., Table 4.1). Here, migration of C1, V1, M1, and Syllable was examined. Material included five sets of pseudoword pairs, corresponding to five sets of target words: *hairu* and *gendo*, *shukko* and *zange*, *shusse* and *maido*, *dappi* and *moeru*, and *shukka* and *soegi*. Thus, one set included CVVCV and CVNCV sequences, one included CVQCV and CVNCV sequences, and three sets included CVQCV and CVVCV sequences.

In terms of prosodic structure (which is important, as pitch accent is lexically contrastive in Japanese – e.g., Haraguchi, 1999; Shibatani, 1990), all words (and hence all pseudoword stimuli) were LHH (with syllables bearing a pitch accent, labeled high, *H*, in contrast to syllables labeled low, *L*), except for the *hairu* and *gendo* set, which was HLL.¹ The accent patterns of all the stimuli were verified in a Japanese accent dictionary (NHK, 1998).

As illustrated in Table 4.1, for each examined constituent, all target

constituents were present but distributed between the two stimuli of the dichotic stimulus pair in the experimental trials. This allows erroneous perception (i.e., FAs) of the target word by information blending between the two ears. However, FAs may occur due to mere misperception rather than migration of some phonological constituent.

To control for this possibility, control trials were employed. These differed from the experimental trials in terms of only one stimulus of the dichotic pair. For C1 and V1, only one type of control trial was used. Indeed, for these constituents it was easy to decide which stimulus of the experimental pair was perceptually closer to the target: as illustrated in Table 4.1, one dichotic pseudoword shared all phonemes with the target except one (C1 or V1), whereas the other differed from the target in terms of all phonemes except one (C1 or V1). Therefore, for C1 and V1, the control trials included the stimulus that, in the experimental pair, was perceptually closer to the target, and the other stimulus of the experimental pair was changed minimally to lack the critical phoneme that would have allowed a migration error. For M1 and Syllable, it was difficult to decide a priori which of the two experimental stimuli was perceptually closer to the target. Thus participants were presented two different control pairs for M1 and Syllable: one shared with the target (and hence with the experimental pair) the initial portion, the other shared the final portion (see Table 4.1).

Thus, there were 10 different types of target-absent trials for each target. Each type of target-absent trial was presented six times: three times for each ear assignment pattern (e.g., for the C1 experimental trials, *gairu* was presented to the left ear and *hendo* to the right ear, or vice versa; see Table 4.1).

There were also 14 different types of target-present trials for each target. The target-present trials were constructed by replacing either of the original stimuli of the dichotic pair with the target (e.g., for C1 and the target *gendo*, the target-present trials paired *gendo* with *gairu*, *hendo*, or *kairu*). Each type of target-present trial was presented twice, counterbalanced for ear assignments.

This yielded 880 trials, which were preceded by 60 training trials

constructed from different sets (with the words *gunte* and *kaori*, *denju* and *taiho*, *denbu* and *kappo*, and *dango* and *mekki* as targets) paired according to the same principles. The whole experiment included 180 additional trials that were not analyzed; they were intended to test migration of the second mora, but this led to serious interpretation problems given the use of geminate stop consonants (realized as silence, cf. Fujimura, 2007; Fukumori, 2010; Muraki & Nakaoka, 1990; Saito, 2003; Sato, 2001; Takayama, 2011) in four of the five sets.

The pseudoword stimuli and the target words were pronounced naturally by a native speaker of Japanese and were recorded in an anechoic room. All stimuli were digitized at 16-bit and 32 KHz sampling rate on a computerized speech workstation (*DigiDesign's Sound Tools II*TM running on a Macintosh II FX). Each mono file corresponding to an individual stimulus (pseudoword or target word) was then normalized (scaling all values upward so that the peak is at the maximum allowable value of the dynamic envelope – hence achieving optimal dynamic range).

In each dichotic pair, the stimuli were synchronized as much as possible at both the onset and offset of the critical constituents. a perceptual criterion was used: first, for each isolated (mono) stimulus (e.g., *geiru*, *hando*), a native speaker of Japanese (the author) told the research assistant who was operating the digital audio workstation when she thought she heard the onset and offset of the critical constituent (e.g., the beginning/end of M1). After that, to synchronize the dichotic stimuli, when necessary, the duration difference between the two sequences (e.g., that between *ge* and *ha* in the *geiru-hando* pair or between *gen* and *hai* in the *genru-haido* pair) was calculated and then equalized both by cutting the appropriate number of cycles in the steady state part of the first and/or final vowel of the longer sequence and by multiplying one cycle of the most stable part of the first and/or final vowel of the shorter sequence as much as necessary (see Figure 4.2).

Thus, the co-articulatory indices between consonant and vowel as well as between morae and syllables were not affected by the signal manipulations. Finally, the author perceptually checked the result of this

synchronization in order to confirm the naturalness (including rhythmic normality) of each stimulus. The average duration of the 225 different stimuli we measured (excluding repetitions) using *Praat* software (Boersma, 2001; Boersma & Weenik, 2013) was 812 ms (SD: 54 ms).

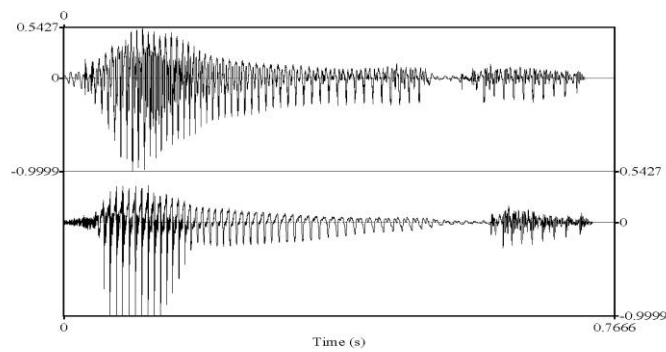


Figure 4.2. Illustration of the waveforms of the stimuli of a dichotic pair used in the experimental M1 condition of Experiment 1 for the detection of the target word *gendo* or *hairu*. The waveforms corresponding to *geiru* and *hando* are on the top/bottom, respectively. For both, the numbers on the left (for *geiru*) or the right (for *hando*) of the Y axis correspond to the amplitude; the duration of both stimuli (X axis) was 737 ms, with a total file duration (final silence included) of 767 ms. The measurements and figure were made using *Praat* software (Boersma, 2001; Boersma & Weenik, 2013).

Procedure: Trials were presented pseudorandomly. The stimuli were played back through headphones by means of a digital recorder at approximately 70 dB SPL with a 2-s silent interval between successive trials. Ear assignment of headphones was counterbalanced between participants. As illustrated in Figure 4.1, the participants were given a booklet in which the written targets were listed for each trial. For one group of participants, targets were presented only in hiragana, while for the other group, targets were presented either in kanji alone or in kanji plus a hiragana suffix (okurigana). In the first case, the three written symbols corresponded to the three morae of the word (e.g., げんど, /ge.N.do/; だっぴ, /da.Q.pi/; はいる, /ha.J.ru/). In the second case, the kanji and hiragana corresponded to syllables for most of the words, whether they were written with two kanji characters, as in *gendo* 限度 (/geN/ in kanji + /do/ in kanji) and *dappi* 脱皮 (/daQ/ in kanji + /pi/ in kanji), or with one kanji and one kana character, as in *hairu* 入る (/haJ/ in kanji + /ru/ in hiragana)².

Participants were told that during each trial, they would hear two speech stimuli that could be meaningful or meaningless and were asked to pay attention to both items (i.e., to both ears) in order to decide whether or not the target indicated in the booklet for each particular trial had been presented; they answered by drawing either a plus or a minus sign, respectively. The experimental session lasted about 1 hour.

4.3.2 Results

The d' scores were calculated for each subject from the hit (correct detection of the target on target-present trials) and FA (incorrect detection of the target on target-absent trials) percentages³. The average values of all these scores are presented in Table 4.2, separated according to experimental vs. control trials, speech constituent, and target presentation (hiragana vs. kanji).

An ANOVA performed on d' scores⁴ took into account constituent and trial type (experimental vs. control trials) as within-subjects variables plus target script (hiragana vs. kanji) and country of test administration

(Belgium vs. Japan) as between-subjects variables. It showed significant effects of constituent, $F(3, 156) = 127.42, p < .0001$, and trial type, $F(1, 52) = 262.75, p < .0001$. There were no significant effects of target script or country of test, both $F_s < 1$, and these variables did not interact with the other ones, all $F_s \leq 1$. There was however a significant interaction between constituent and trial type, $F(3, 156) = 57.95, p < .0001$, which is illustrated in Figure 4.3. This interaction showed that, although the differences between the experimental and control trials were significant for all speech constituents (Bonferroni-corrected post-hoc tests: $F_s(1, 156) = 30.63, 17.37, 401.76$, and 223.64 for C1, V1, M1, and Syllable, respectively, all $ps < .001$), this difference was larger for M1 than for any other constituent (vs. C1: $F(1, 156) = 105.26$; vs. V1: $F(1, 156) = 126.03$, both $ps < .001$), including Syllable, $F(1, 156) = 12.95, p < .005$. Syllable also yielded a larger difference between experimental and control trials than did C1 and V1, $F(1, 156) = 44.37$ and $= 56.18$, respectively, both $ps < .001$.

Table 4.2. Average percentages of false alarms on target-absent trials and of correct detections on target-present trials, plus average d' scores for each speech constituent tested on the Japanese-speaking participants examined in Experiment 1, separately for the targets presented in hiragana vs. kanji, and overall.

Standard deviations in brackets.

Constituent	Target script	Trial Type					
		Target-absent		Target-present		d' scores	
		Experimental	Control	Experimental	Control	Experimental	Control
C1	kana	35.54 [21.01]	23.69 [14.38]	92.68 [4.35]	92.23 [4.48]	1.97 [0.79]	2.31 [0.72]
	kanji	27.98 [15.73]	18.39 [11.45]	91.25 [5.20]	90.71 [4.99]	2.11 [0.66]	2.39 [0.58]
	average	31.76 [18.78]	21.04 [13.16]	91.96 [4.81]	91.47 [4.76]	2.04 [0.73]	2.35 [0.65]
V1	kana	11.37 [13.60]	8.75 [11.70]	91.16 [6.75]	92.14 [6.00]	2.93 [0.78]	3.14 [0.78]
	kanji	7.32 [7.91]	5.24 [7.15]	87.05 [8.39]	88.57 [7.05]	2.87 [0.74]	3.12 [0.80]
	average	9.35 [11.21]	6.99 [9.77]	89.11 [7.82]	90.36 [6.73]	2.90 [0.76]	3.13 [0.78]
M1	kana	14.52 [14.98]	1.96 [2.83]	90.71 [6.66]	92.90 [4.93]	2.70 [0.80]	3.82 [0.65]
	kanji	10.71 [10.08]	0.89 [1.62]	90.89 [5.58]	92.72 [4.21]	2.82 [0.71]	3.95 [0.47]
	average	12.62 [12.80]	1.43 [2.34]	90.80 [6.09]	92.81 [4.54]	2.76 [0.75]	3.88 [0.56]
Syllable	kana	11.43 [13.51]	1.99 [2.53]	91.70 [5.93]	92.63 [5.26]	2.97 [0.90]	3.78 [0.62]
	kanji	7.08 [7.75]	0.89 [1.36]	89.46 [7.86]	90.98 [6.42]	3.021 [0.63]	3.89 [0.61]
	average	9.26 [11.13]	1.44 [2.09]	90.58 [6.99]	91.81 [5.88]	2.99 [0.77]	3.83 [0.61]

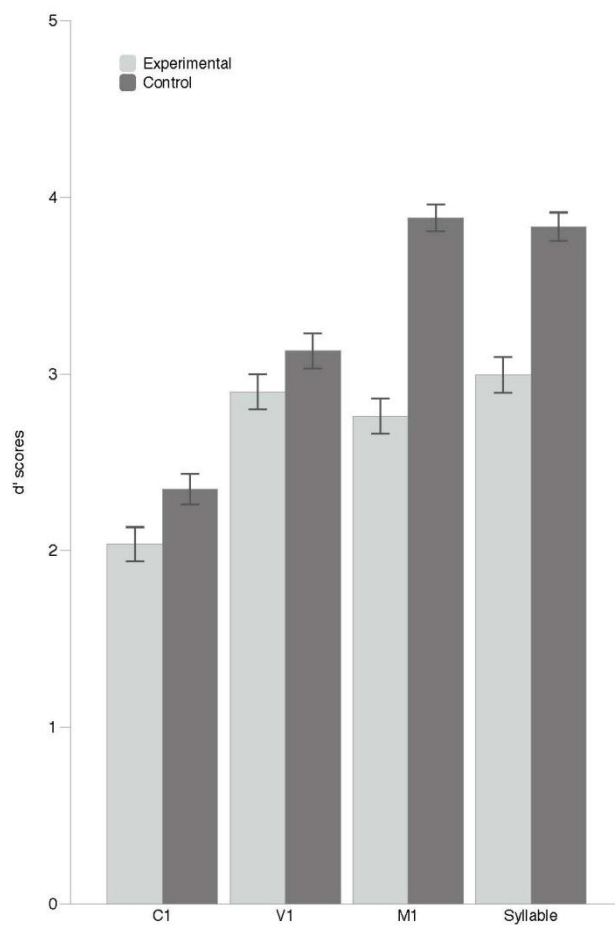


Figure 4.3. Average values of d' scores, separately for each speech constituent and trial type (experimental vs. control), observed on the Japanese-speaking participants examined in Experiment 1.

Error bars represent standard errors.

4.3.3 Discussion

The results of the present experiment suggest that Japanese speech segmentation relies on all the manipulated segments, but primarily on both syllables and morae, and that the morae provide the basis of the preferred segmentation strategy. In addition, the facts that we observed migrations of consonantal phonemes for which there exist no isolated written signs and that we did not observe any influence of target script (i.e., of the fact that the targets were presented in either kana or kanji) suggest that the speech segmentation procedure was not influenced by orthography. Before examining this idea at more length in Experiment 3, in the next experiment, we checked whether the response pattern observed here is specific to Japanese listeners.

4.4 Experiment 2: French Listeners

To examine whether the response pattern observed in Experiment 1 among Japanese speakers reflects native language segmentation processes rather than material-dependent strategies, in Experiment 2, we presented the same material as that used in Experiment 1 to French-speaking listeners. Although no effect of target script was observed in Experiment 1, we preferred to present targets orally rather than alphabetically. Indeed, for native speakers of French, an alphabetic transcription of the target would have been ambiguous or approximate for many of the target words, especially those including a nasal mora or a geminate like *gendo* and *shukko*, respectively.

4.4.1 Method

Participants: Thirty-six students of the Université Libre de Bruxelles, Belgium, participated in the experiment for course credit. All were native speakers of French without any knowledge of Japanese; although most had

some notions of English and/or Dutch, none was a fluent bilingual. Nine were discarded because they missed the target quite often (i.e., correct detection scores < 82%, which was the lowest score observed among the Japanese participants in Experiment 1). The final sample thus included 27 participants (aged 18–23 yrs; average: 19 yrs; 22 women).

Material and procedure: The material and procedure were exactly the same as in Experiment 1, except that the prerecorded targets (pronounced by the same Japanese speaker as the one who recorded the stimuli) were presented orally before each pseudoword pair. Specifically, in each trial, participants first heard the target word, then a short (100-ms) warning beep, and – after a 500-ms silence – the dichotic pair, with a 3-s silent interval between successive trials. Since this procedure took more time than with written target presentations, the experiment was run in 2 sessions of about 40 min each, which were temporally separated by at least 1 day.

4.4.2 Results

As in Experiment 1, we analyzed the d' scores computed on the basis of the FA and hit values, the average values of which are presented in Table 4.3.

An ANOVA on d' scores (the average values of which are presented in Figure 4.4) showed a significant interaction between constituent and trial type (experimental vs. control), $F(3, 78) = 10.74$, $p < .0001$, in addition to main effects of constituent, $F(3, 78) = 104.56$, and trial type, $F(1, 26) = 70.78$, both $ps < .0001$. Bonferroni-corrected post-hoc tests showed that all constituents except V1, $F(1, 78) = 4.06$, $p > .10$, led to significant effects of trial type, $F_s(1, 78) = 27.42$, 99.15, and 40.59 for C1, M1, and Syllable, respectively, all $ps < .001$. In addition, the difference between the experimental and control trials was smaller for V1 than for M1 or Syllable, $F_s(1, 78) = 31.74$ and 9.49, $p < .05$, and smaller for C1 than for M1, $F(1, 78) = 11.14$, $p < .025$. No significant difference was observed between M1 and Syllable, $F(1, 78) = 6.43$, $p > .10$.

Table 4.3. Average percentages of false alarms on target-absent trials and of correct detections on target-present trials for each speech constituent tested on the French-speaking participants examined in Experiment 2. Standard deviations in brackets.

	Trial Type			
	Target-absent		Target-present	
	Experimental	Control	Experimental	Control
C1	24.07 [13.13]	15.62 [9.17]	88.43 [7.21]	89.72 [6.29]
V1	4.94 [7.08]	3.27 [5.36]	87.59 [6.26]	87.87 [5.53]
M1	5.00 [6.75]	0.62 [1.02]	89.35 [5.95]	91.20 [5.66]
Syllable	3.33 [4.83]	0.40 [0.74]	91.02 [5.34]	89.95 [5.56]

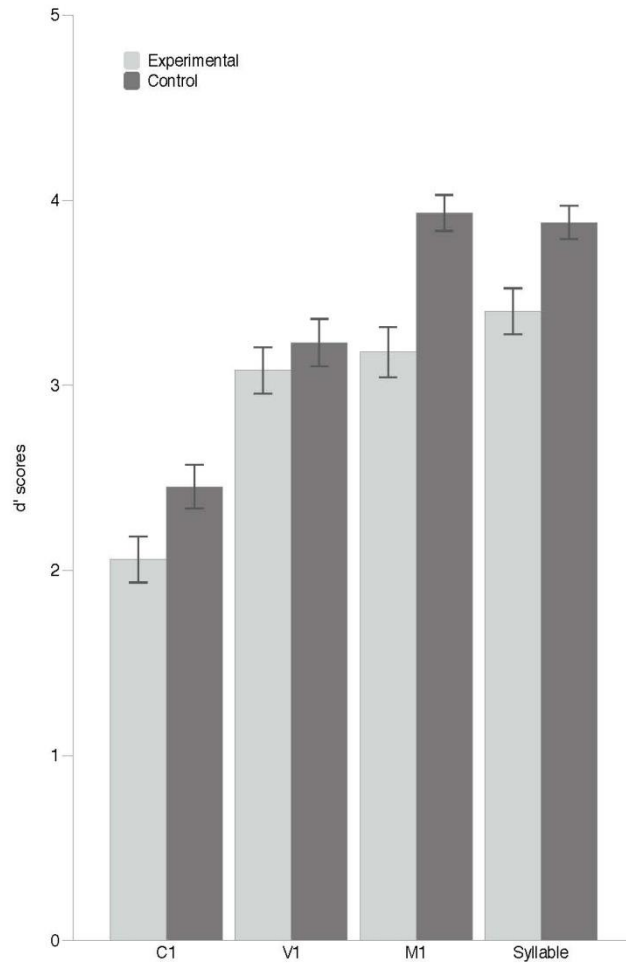


Figure 4.4. Average values of d' scores, separately for each speech constituent and trial type (experimental vs. control), observed on the French-speaking participants examined in Experiment 2.

Error bars represent standard errors.

4.4.3 Discussion, cross-experiment comparison, and control metaphonological experiment

Contrary to the pattern observed among Japanese listeners in Experiment 1, French-speaking participants did not display superiority of M1 over the Syllable.

Cross-experiment analysis confirmed that French-speaking listeners did not segment the Japanese material used here in the same way as the Japanese speakers examined in Experiment 1 did. Indeed, the ANOVA showed a significant interaction between constituent, trial type, and participant's native language (French vs. Japanese), $F(3, 243) = 5.47$, $p = .001$. While the French-speaking participants did not differ from the Japanese speakers for C1 and V1 (trial type $\times$ group interaction: both F s < 1), their effects of trial type were significantly smaller for M1 and Syllable compared with those of the Japanese participants in Experiment 1 (trial type $\times$ group interaction: $F(1, 81) = 10.99$ and 8.51 , respectively, both p s $< .005$).

Thus, the superiority of mora- over syllable-based migration observed in Experiment 1 seems specific to native speakers of Japanese. However, it remains to be understood why the French-speaking listeners showed both morae and syllable migrations. Indeed, their native language phonology would lead us to expect them to use mainly syllables, not morae, both because morae are not relevant in their native language and because some morae (particularly /Q/ in CVQCV items, e.g., *shu-s-se*), would be difficult to perceive, as speakers of non-geminate languages like French have difficulties perceiving the contrast between Japanese geminate and singleton consonants (Hardison & Motohashi Saigo, 2010).

A possible explanation of the morae migrations observed in French-speaking listeners may be that they interpreted the other morae (e.g., those including a single vowel) as syllables. This possibility was checked by a control metaphonological experiment on 10 naïve native speakers of French (aged 17–29 yrs, mean age: 24 yrs, 5 women) who did not participate in the previously described experiments and who had no

knowledge of Japanese at all. They were asked to decide how many “salient units” they heard in the 50 spoken sequences that had served as stimuli or targets in Experiment 2. No examples were provided so as not to bias their answers. They were asked to answer without thinking too much, following their intuition, and they were told that there were neither “good” nor “wrong” answers. Overall, the presence of two salient units was reported in the majority (58%) of responses. However, the overall rate of responses reporting the presence of three salient units (which, for native speakers of Japanese, would correspond to morae) was not negligible (overall: 38%). In fact, the latter responses were dominant (79%) for the CVVCV sequences like *hairu*, *moeru*, *maido*, *soegi*, and derived stimuli like *shuido*, *soeka*, or *haido*. On the contrary, the CVNVCV and CVQCV sequences led to a large majority of responses reporting the presence of two salient units (72% and 95.5%, respectively). An ANOVA on the arcsine transform of the proportion of “three” responses showed a significant effect of sequence structure (CVVCV, CVNVCV, or CVQCV), $F(2, 18) = 52.23$, $p < .0001$, with all post-hoc contrasts confirming significance at $p < .005$ (Bonferroni-corrected). These results show that, with CVVCV items (e.g., *hairu*), French speakers consider the second mora (a vowel) as a syllable, which is not the case when the second mora is /N/ or a geminate consonant.

In the migration experiment, most dichotic stimuli (except for the *shukko-zange* set of targets) were pairs mixing either CVVCV and CVNVCV or CVVCV and CVQCV sequences (see e.g., Table 5); the perceptual result of such mixing was probably quite ambiguous for the French-speaking listeners who had been tested in Experiment 2. Indeed, the presentation of at least one CVVCV stimulus out of a pair or a CVVCV target may have enhanced their rate of presumed moraic migrations, which in fact were syllable-based.

The results of the metaphonological test on French-speaking listeners may also account for why the French speakers showed fewer syllable migrations than the Japanese speakers from Experiment 1. Indeed, in CVVCV sequences like *hairu*, syllables like *hai* and *ru* do not seem to have

psychological reality for French natives, as they consider these stimuli as including three salient units (presumably syllables): *ha*, *i*, and *ru*.

Another possibility is that, in Japanese listeners, the syllable migrations reflect the influence of kanji representations, which may be evoked mentally by Japanese participants, whatever the target script. This possibility was examined in the next experiment.

4.5 Experiment 3: Are Migrations in Japanese Listeners Affected by the (Un)availability of Kanji Representations?

The aim of the third experiment was to ascertain that the migrations of syllables were not affected by the availability, in Japanese readers, of kanji written representations. Indeed, although any influence of the target script (hiragana vs. kanji) in Experiment 1 was not observed, Japanese participants may have activated kanji representations regardless of the actual target script they were presented with. Although all the Japanese target words used in Experiment 1 can be transcribed by means of kana characters only, they are usually transcribed by means of two kanji characters or a kanji character plus a hiragana suffix. If the speech segmentation procedure were influenced by written representations, this would have favored syllabic migrations in Experiment 1, as syllables corresponded to kanji in that experiment's material.

To test this idea, in this experiment, we presented as targets only katakana loanwords for which no kanji representations exist; we did this to exclude the possibility that syllabic migrations could result from the activation of mental representations of kanji. If the speech segmentation procedure were influenced by orthographic representations, and if the kanji representations were responsible for the high rate of syllabic migrations observed in Experiment 1, we would observe with the present material far fewer migrations of syllables than in Experiment 1; here, this would lead to a more pronounced advantage of M1 over Syllable.

Furthermore, the present material was designed to allow examination of migrations of the second mora (henceforth, *M2*) in addition to those of the other speech constituents already examined in Experiment 1. This was not possible in Experiment 1, because its targets were many words (in four of the five target sets) that included geminate stop consonants. In that case, /Q/, sokuon, the first member of the geminate, is realized as silence (Fukumori, 2010; Fujimura, 2007; Muraki & Nakaoka, 1990; Saito, 2003; Sato, 2001; Takayama, 2011). This would lead to serious interpretation problems regarding *M2* migration errors. In the present experiment, such a problem was avoided by using only sequences with CVJCV-CVNVCV structure (e.g., *raisu-konte*, /raJsu/, /koNte/). In this way, we could examine the role of the “special morae” /N/ (hatsuon) and /J/, [i], the second member of a diphthong. As the other special morae /R/ and /Q/, these constitute the latter part of a heavy syllable and probably do not have the same phonological independence as CV or V morae (e.g., Kubozono, 1999b; Tanaka, 2000, 2008; Ujihira, 1996). Finally, the /oe/ sequence used in the former experiments (as in *moeru* and *soegi*) was also excluded from the target words here, as it is usually not admitted as a syllable in Japanese phonology (Fujimura, 2007; Kasuya & Sato, 1990; Kubozono & Ôta, 1998; Saito, 2003; Shirota, 1993).

4.5.1 Method

Participants: Twenty-eight students at Keio University’s Shonan Fujisawa Campus (SFC) participated in the experiment (aged 20–26 yrs; average: 22 yrs; 10 women). They were all native speakers of Japanese, 68% of them being from the Tokyo area. They were paid for their participation.

Material and procedure: The material and procedure were the same as those employed in Experiment 1, except for the following facts. All the targets were Japanese loanwords (all common nouns except for a proper noun, *Dante*), and all can only be written in Japanese using katakana. Three sets of target words (and pseudoword) pairs were used: *raisu* and *konte*, *toire*

and *pantsu*, and *boiru* and *Dante*. Thus, all the sequences had a CVJCV-CVNCV structure, and all had the HLL pitch pattern (see Table 4.4).

As in Experiment 1, all the target words (and derived pseudowords) were intended to include three morae and two syllables, as they were either CVJCV or CVNCV sequences. To further check that native speakers of Japanese do consider the target words as including three morae and two syllables, as in Experiment 1, a metaphonological task was used. The target words were presented to independent, naïve native speakers of Japanese (eight participants, aged 20–42 yrs; average: 30 yrs; 6 women) and asked them to indicate where they would segment the heard word. The proportion of moraic- (or kana-) based responses was slightly higher than in Experiment 1 (17%); the proportion of “two” responses was lowered to 71%⁵. Thus, as in Experiment 1, there is a preference for a syllabic mode of explicit, intentional, segmentation that — although surprising — is in agreement with the idea that the stimuli had a disyllabic structure.

The experimental and control trials were constructed as in Experiment 1, except that M2 experimental and control trials were added for both target-absent (e.g., for the target *raisu*, *ransu* paired with *koite* as an experimental trial and *ransu* paired with *koete* as a control trial; see Table 4.4) and target-present pairs; this yielded a total of 636 trials. These were preceded by 30 training trials, which were constructed from different sets of words (*sairo* and *bentsu*, *daisu* and *kompe*, and *faito* and *renzu*) and pseudowords paired according to the same principles. The average duration of the 162 different stimuli measured (excluding repetitions) by *Praat* software (Boersma, 2001; Boersma & Weenik, 2013) was 574 ms (SD: 41 ms). The procedure was the same as that of Experiment 1, except that in the response booklet, the targets were written in katakana. The experimental session lasted about 35 min.

Table 4.4. Example of target-absent trials designed for the pair of target words *raisu* (/raJsu/) and *konte* (/koNte/). Letter case signals the distribution of information between the ears: lower case corresponds to parts of the target *raisu*, upper case corresponds to parts of the target *konte*.

Bold characters signal the parts of the experimental trials that have been changed in the control trials. M1('), M2('), and M3(') are the first, second, and third mora, respectively. C: consonant; V: vowel.

Tested constituent	Target	Trial type	Dichotic pair											
			Pseudoword presented to one ear						Pseudoword presented to the other ear					
			M1		M2	M3		M1'		M2'	M3'			
			C1	V1			C2	V3	C1'	V1'			C3'	V2'
C1	raisu or konte	Experimental	K	a	i	s	u		r	O	N	T	E	
	raisu	Control	K	a	i	s	u		d	O	N	T	E	
	konte	Control	P	a	i	s	u		r	O	N	T	E	
V1	raisu or konte	Experimental	r	O	i	s	u		K	a	N	T	E	
	raisu	Control	r	O	i	s	u		K	e	N	T	E	
	konte	Control	r	E	i	s	u		K	a	N	T	E	
M1	raisu or konte	Experimental	K	O	i	s	u		r	a	N	T	E	
	raisu	Control (on initial)	K	O	e	sh	o		r	a	N	T	E	
	raisu	Control (on final)	K	O	i	s	u		m	e	N	T	E	
	konte	Control (on initial)	K	O	i	s	u		r	a	U	P	A	
	konte	Control (on final)	G	U	i	s	u		r	a	N	T	E	
M2	raisu or konte	Experimental	r	a	N	s	u		K	O	i	T	E	
	raisu	Control	r	a	N	s	u		K	O	e	T	E	
	konte	Control	r	a	U	s	u		K	O	i	T	E	
Syllable	raisu or konte	Experimental	K	O	N	s	u		r	a	i	T	E	
	raisu	Control (on initial)	K	O	N	sh	o		r	a	i	T	E	
	raisu	Control (on final)	K	O	N	s	u		d	e	o	T	E	
	konte	Control (on initial)	K	O	N	s	u		r	a	i	P	A	
	konte	Control (on final)	G	U	A	s	u		r	a	i	T	E	

4.5.2 Results

As in the former experiments, we analyzed the d' scores computed on the bases of FAs and hits, the average values of which are presented in Table 4.5.

Table 4.5. Average percentages of false alarms on target-absent trials and of correct detections on target-present trials for each speech constituent tested on the Japanese-speaking participants examined in Experiment 3. Standard deviations in brackets.

	Trial Type			
	Target-absent		Target-present	
	Experimental	Control	Experimental	Control
C1	39.48 [12.36]	26.98 [12.65]	88.24 [6.99]	87.05 [8.73]
V1	4.56 [7.61]	2.18 [4.68]	88.84 [9.00]	93.30 [6.64]
M1	7.14 [7.35]	0.74 [1.53]	89.58 [7.48]	88.24 [6.39]
M2	9.13 [8.88]	7.54 [7.89]	84.67 [9.95]	85.27 [8.06]
Syllable	7.04 [9.69]	0.59 [1.09]	90.48 [8.55]	89.43 [6.29]

The ANOVA on d' scores (see average values in Figure 4.5) took trial type and constituent as factors. It showed significant effects of constituent, $F(4, 108) = 94.72$, $p < .0001$, and trial type, $F(1, 27) = 86.46$, $p < .0001$, as well as a significant constituent $\times$ trial type interaction, $F(4, 108) = 7.94$,

$p < .0001$. Bonferroni-corrected post-hoc tests showed that the difference between the experimental and control trials was significant for all speech constituents, $F_s(1, 108) = 17.69, 28.09, 69.08$, and 61.88 for C1, V1, M1, and Syllable, respectively, all $ps < .001$, except that for M2, $F(1, 108) = 2.04$, $p > .10$. In addition, the difference in scores between the experimental and control trials was larger for M1 than for C1 and M2, $F_s(1, 108) = 8.43$, $p < .05$, and $= 23.68$, $p < .001$, respectively. Syllable also generated a larger difference between the experimental and control trials than did M2, $F(1, 108) = 20.72$, $p < .001$, but the results for Syllable did not differ from those for M1, $F < 1$. No other comparison reached significance (all $ps > .10$), although V1 tended to elicit a larger difference between experimental and control trials than did M2, $F(1, 108) = 7.49$, $p = .07$.

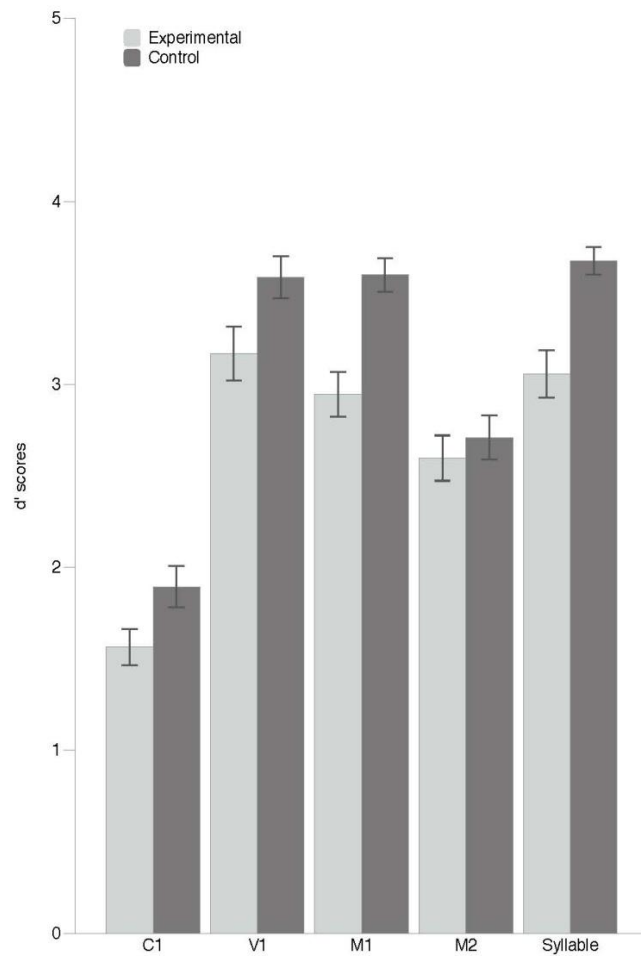


Figure 4.5. Average values of d' scores, separately for each speech constituent and trial type (experimental vs. control), observed on the Japanese-speaking participants examined in Experiment 3.

Error bars represent standard errors.

4.5.3 Discussion and Cross-experiment Comparison

Similarly to the pattern of results observed in Experiment 1, in the present experiment, migrations were observed at various sublexical levels, including C1. However, contrary to the situation in Experiment 1, here, M1 did not migrate more often than Syllable.

In order to compare more directly the result patterns of the two experiments, a cross-experiment ANOVA was run including the data of the present experiment and those collected on the participants of Experiment 1 who were also tested in Japan⁶. discarding the M2 data (which have no counterpart in Experiment 1). In particular, we wanted to check whether migrations were affected by orthography. If this were the case, in the present experiment we would observe more migrations of M1 and less migrations of syllables compared with Experiment 1, as the targets used here corresponded to words that can only be written (and which were presented) in katakana, which correspond to morae.

No overall effect of group (i.e., material), $F < 1$, was observed but a significant interaction between group, constituent, and trial type, $F(3, 162) = 4.90$, $p < .005$. Yet, this interaction reflects an effect opposite to the prediction. Indeed, a significant group $\times$ trial type interaction was observed for M1, $F(1,54) = 6.04$, $p < .025$, with a larger trial type effect observed in

Experiment 1 compared with the present experiment. The group $\times$ trial type interaction also was marginally significant for V1, $F(1,54) = 3.79$, $p = .057$, with a somewhat *larger* trial type effect observed in the present experiment than in Experiment 1. No other interaction between group and

trial type was significant, even for Syllable, $F(1,54) = 2.05$, $p > .10$ (C1: $F < 1$).

Thus, if there were an effect of material on the response pattern, it was an unexpected one. In particular, contrary to the pattern of results observed in Experiment 1, here, M1 did not migrate more than Syllable and actually migrated significantly less than it did in Experiment 1. This does not support the idea that migrations were affected by orthography. Thus, the

availability of kanji representations in Japanese participants cannot account for the numerous syllabic migrations observed both here and in Experiment 1; neither can the mora-based segmentation procedure, as revealed by migration errors, reflect only orthographic representations.

4.6 General Discussion

Speech segmentation was examined in Japanese, a language in which the mora has been proposed as central to this process (e.g., Otake et al., 1993). Specifically, the aim was to compare the involvement of several speech constituents in a task that does not require listeners to analyze the input into sublexical constituents intentionally. Indeed, former evidence — mostly gathered through the fragment detection task (Mehler et al., 1981) — may have been influenced by orthographic knowledge and/or associated metalinguistic representations (e.g., Inagaki et al., 2000). In addition, several pieces of evidence show that phonemes and syllables may also play important roles in Japanese (Cutler & Otake, 2002; Kawahara, 2007; Kawahara & Shinohara, 2009a; b; Kubozono, 1998, 2003; Tamaoka & Taft, 1994; Tanaka, 1999, 2008).

We studied a perceptual illusion, namely migration errors between dichotic stimuli (Kolinsky & Morais, 1996; Kolinsky et al., 1995) in a target detection task: illusory perceptions (or slips-of-the-ear) of the target word by migration of some sublexical constituent (phonemes, morae, or syllables) from one stimulus of the dichotic pair to the other would reveal the implicit representations of the perceptual, preattentive (cf. Treisman & Paterson, 1984) constituents of speech.

In order to control for the impact of literacy-associated knowledge on segmentation processes, in Experiment 1, we presented native speakers of Japanese with target words that could be written and were actually presented in either hiragana and/or kanji (precisely, kanji or a kanji-hiragana combination), in which written symbols correspond to either morae or syllables, respectively. No effect of target script was observed.

Across both target script conditions, the results suggested that Japanese speech segmentation relies on several sublexical constituents, including consonantal phonemes (the first consonant, C1) for which there exist no isolated written characters. However, both Syllable and the first mora (M1) seem more important, with the mora providing the basis of the preferred segmentation strategy.

To examine whether this response pattern is specific to Japanese listeners, in Experiment 2, we presented the same material as that used in Experiment 1 to French-speaking listeners, with targets presented orally. Native speakers of French did not segment the Japanese material used here in the same way as the Japanese speakers of Experiment 1 did. In particular, they produced fewer migrations for M1 and Syllable compared to the native speakers of Japanese, and, contrary to the latter, did not present more M1 than syllable migrations. This indicates that the results of Experiment 1 reflect Japanese native language segmentation processes rather than material-dependent strategies.

In addition, a post-hoc metaphonological test suggested that French speakers interpreted the Japanese morae as syllables, which may explain why they produced both morae and syllable migrations. The metaphonological results may also help to explain why the French speakers of Experiment 2 showed fewer Syllable migrations than the Japanese speakers who participated in Experiment 1 did: in CVVCV sequences like *hairu*, syllables like *hai* and *ru* do not seem to have psychological reality for native speakers of French, as they consider that these stimuli include three salient (probably syllabic) units: *ha*, *i*, and *ru*.

Experiment 3 explored an alternative explanation of the fact that Japanese listeners obtained so many syllable migrations in Experiment 1, namely, the possible influence of activation of kanji mental representations. Although we did not observe any influence of the target script in Experiment 1, we acknowledge that presenting the targets of Experiment 1 using only hiragana (kana) characters is somewhat unusual, as these words are typically written with at least one kanji character. Hence, Japanese participants might resort to mental representations of kanji regardless of the

actual target script. If orthographic representations influenced the speech segmentation procedure, this would have favored syllabic migrations upon exposure to the material used in Experiment 1, in which one kanji character corresponded to one syllable.

To test this idea, the targets in Experiment 3 were only katakana words (i.e., loanwords), for which there exist no kanji representations at all. Furthermore, the material was designed to allow the observation of migration of the second mora (M2) in addition to that of other speech constituents already examined in the former experiments. Specifically, migrations of “special morae” (/N/, and /J/, [i]) were examined. Indeed, these constitute the latter part of a heavy syllable and probably do not have the same phonological independence as V or CV morae (e.g., Kubozono, 1999b; Tanaka, 2000, 2008; Ujihira, 1996).

As in Experiment 1, C1 migrations were observed in Japanese listeners, indicating the intervention of orthography-independent sublexical constituents. In addition, compared with Experiment 1, the Japanese listeners of Experiment 3 did not show more M1 than Syllable migrations: both speech constituents led to similar levels of illusory percepts. This unexpected result confirms that the availability of kanji representations to Japanese participants cannot account for the numerous syllabic migrations observed in Experiments 1 and 3. As revealed by migration errors, the mora-based segmentation procedure also cannot reflect orthographic (or associated metalinguistic) representations. Had this been the case, we would have observed fewer Syllable and more M1 migrations in Experiment 3 compared with Experiment 1, as Experiment 3’s targets corresponded to words that can only be written (and which were presented) in katakana.

Nevertheless, the fact that we did not observe more M1 than Syllable migrations in Experiment 3 questions the generalizability of Experiment 1’s results. In Experiment 1, we used several items (four out of the five target sets) that included geminate stop consonants, which was not the case in Experiment 3. This may have been problematic, as the first member of a geminate stop consonant is realized as silence, as is the latter part (/e/) of the /oe/ diphthong used with the targets *moeru* and *soegi* in Experiment 1.

Further, in Experiment 3, only HLL stimuli were used, and in all dichotic pairs CVJCV and CVNCV were mixed; this was the case for only one set used in Experiment 1 (i.e., the targets *hairu* and *gendo*). The material of Experiment 3 was thus more systematic and refined than that of Experiment 1. Further experiments are needed to identify which of these factors influenced the response pattern.

Whatever this factor may be, even the more controlled material of Experiment 3 led to numerous syllabic migrations, as did the material of Experiment 1. As the migration task did not require participants to analyze the stimuli intentionally, it is highly likely that it tapped earlier, less-conscious processing than, for example, the fragment detection task used in previous work (e.g., Otake et al., 1993; Inagaki et al., 2000). Thus, although the mora may be the main rhythmic and/or segmentation unit in the latter kind of task in Japanese, the processing units could differ at an earlier perceptual, preattentive, level. The present migration results are in fact in agreement with other data. For example, Cutler and Otake (2002, p. 312) stated that their similarity and lexical decision findings “do not support the suggestion that the mora functions as an early structural unit in Japanese listener’s recognition of spoken words” and argued that “although morae are more accessible to Japanese listeners’ awareness than phonemes are, morae do not appear to play a direct role in Japanese spoken-word recognition.”

Thus, the evidence seems to indicate that, in addition to morae, both consonantal phonemes and syllables play a role in Japanese early spoken word processing.

This conclusion may seem at odds with the recent claim by Labrune (2006; 2012a; b) that syllables do not exist in Japanese phonology. Following Poser’s (1990) original proposal, this author directly relates the mora to another, higher-level metrical unit, the bimoraic foot, which often (but not always) corresponds to the syllable. Indeed, a bimoraic foot may consist of one or two syllables (e.g., *koto*, meaning “Japanese harp”, includes one foot but two syllables, while *koN*, meaning “navy blue”, includes one foot but only one syllable, cf. Inaba, 1998). According to some authors, bimoraic feet are important in determination of the basic default

accent pattern in Japanese (e.g., Kondo & Arai, 1998; Kubozono, 1999c, 2006; Poser, 1990; Shinohara, 2000; see also Tanaka, 1999, who even considers quadrimoraic units). They also play various roles in the formation of hypocoristics (Mester, 1990; Poser, 1990), musicians' language (Tateishi, 1989), and mimetics (Nasu, 2006; Tsujimura, 2007) and are favored in nursery songs (Machida & Asano, 1962; Ohara, 1994, cited by Kondo, 2006).

According to Itô (1990) and Kubozono (1999b), the most productive word formation pattern in loanword truncation and compound clipping (whereby compound nouns and phrases are shortened) may rely on the bimoraic foot rather than the syllable. Indeed, these processes combine the first two morae of the two original words, whether they correspond to syllables (e.g., in *hebi meta*, meaning “heavy metal”, which becomes *hebi meta*: 2 morae and 2 syllables + 2 morae and 2 syllables) or not (e.g., in *jiiNzu paNtsu*, meaning “jeans pants”, which becomes *jii paN*: 2 morae and 1 syllable + 2 morae and 1 syllable).

Most syllables in the present experiments' materials corresponded to bimoraic feet, as they were heavy syllables composed of one CV mora plus one special mora. Only two targets (*moeru* and *soegi* in Experiment 1) corresponded to bimoraic feet but not Japanese syllables. However, they were matched with CVQCV targets (*moeru* with *dappi* and *soegi* with *shukka*), in which the first bimoraic foot corresponds to a syllable. Hence, whether the bimoraic foot was involved in the migration errors rather than or in addition to the syllable remains an open question. Thus, it will be interesting to examine in future studies whether bimoraic migrations occur with sets of targets in which bimoraic feet and syllables do not coincide (e.g., *soe/gi* and *kao/ri*, with / showing the foot boundary).

A further issue concerns the status of M2. In Experiment 3, we observed that this moraic structure did not give rise to a high rate of migration errors, for reasons yet to be determined. The fact that M2 was always an isolated phoneme (/J/ or /N/), contrary to M1, which was always a CV, cannot by itself account for the entirety of the results, as Japanese listeners produced significant numbers of migrations of the isolated C1 and V1 phonemes in

both Experiments 1 and 3. Similarly, differences in the pitch pattern between the materials do not seem to explain the entirety of the results. In the material of Experiment 3, M2 was a low-pitched, unaccented mora, whereas M1 was the accented, high-pitched mora; yet, M1 migrations were observed in Experiment 1, despite the fact that M1 was unaccented in most of that experiment's target words. Even the combination of these two factors is insufficient to explain the complete results, as in Experiment 1, isolated phonemes (C1 and V1) migrated even when they were part of an unaccented mora.

One possibility is that the synchronization technique which was used to construct the material (see Method section of Experiment 1) resulted in M2 being too misaligned to allow migrations. Another possibility is that the results reflect the particular status of the special morae used as M2 in Experiment 3 (/N/ and /J/). Indeed, although special morae intervene in phenomena like accentuation rules (e.g., in loanwords, if the antepenultimate mora is a special mora, it does not have an accent but yields the accent to the immediately preceding independent mora; Kubozono, 1995), they are considered as “special” precisely because several phonological phenomena suggest that they do not present the same degree of independence as V or CV morae (e.g., Kubozono, 1999b; Tanaka, 2000, 2008; Ujihira, 1996).

Kubozono (1995) states that special morae (e.g., a long vowel /R/ and the final nasal /N/) are called so because they do not have independence (or autonomy) in terms of accent assignment, although each of them functions as an independent unit. Uwano (2003) further declares that the notion of a special mora is based on the premise of the existence of a Japanese syllable (which, as was discussed, might also be a foot): a special mora is one that can constitute one mora but cannot constitute one syllable by itself. It is thus an M2 of a CVM2 only when it comes after a CV mora, in conjunction with which it constitutes one syllable. The results are thus coherent with the idea that special morae are different from V or CV morae, probably because they constitute parts of heavy syllables and/or bimoraic feet.

Thus, the findings from the present experiments using a task that

required no intentional sublexical analysis suggest that the special morae are less salient than other morae in Japanese speech segmentation. More generally, the present data show that not only morae, which are consistent with kana representations, but also syllables (and/or bimoraic feet), which are consistent with kanji representations, and even consonantal phonemes, which have no corresponding written language signs, play important roles in the perception of spoken Japanese by native listeners. The intervention of consonantal phonemes, in particular, shows that there is no effect of orthography at the early, preattentive, stage of spoken-word processing examined here. This is coherent with recent results obtained in the production of Japanese words. Using word reading tasks, Verdonschot, Kiyama, Tamaoka, Kinoshita, La Heij, and Schiller (2011) observed no effect of script (kana vs. romaji, i.e., Romanized Japanese, using alphabetic orthography) on masked priming effects. The observation by those authors of priming effects both when the mora was a syllable and when it was only a part of a syllable is also coherent with the present data.

The experimental method of eliciting illusion needs improvement. Instead of presenting the stimuli in a booklet of answer sheets comprising a list of target words, a “free-answer method” of response – in which participants can provide their own answers for the heard item – could be used to avoid forcing participants to choose among certain words. In fact, Kolinsky et al. (1995) have already employed such a method of “identification task”, asking the participants for “free answers”. However, the result showed no clear difference from the present method of providing the participants with “visual targets presented in the list”. Another possible improvement would be to develop a new computerized system of presenting stimuli one at a time in a computer display, to avoid letting participants glance at the series of target words.

Chapter 5 : Concluding Remarks

5.1 Conclusion

My research examined how phonological representations of the Japanese language differ at two levels in the model of the levels-of-processing approach that was introduced in Chapter 1 (1.4.2) from the attention-oriented viewpoint. The highest level involved “conscious access” to the properties represented in the speech percept, and an “intentional analysis process” at which one pays the most attention (see Chapter 3). The lowest level was “perceptual” processing, the early preattentive level at which attention and consciousness have the least involvement (see Chapter 4).

Chapter 3 examined the phonological representations at the highest level of attention. The findings show that the mora is the most prominent unit, but that the syllable is also possible as a spontaneous reference unit when the mora fails to give a satisfactory result.

Chapter 4 examined phonological representations at the lowest level of attention, the preattentive stage of processing. The evidence suggests that, in addition to morae, both consonantal phonemes and syllables play a role, and thus multiple sublexical units of phonological representation are functionally involved.

This dissertation also investigated how native speakers of the Japanese language are aware of *phoneme* and its role in speech segmentation.

Chapter 3 showed that participants followed a rather indirect path to accomplish the task, indicating that for Japanese native speakers, phonemes *per se* are not easily accessible metaphonological units, while the knowledge of the kana matrix, the so-called gojuonzu, does imply the ability to recognize that a CV kana can be analyzed as a consonant followed by a vowel, which amounts to the ability to perform metalinguistic operations at the phonemic level.

The findings presented in Chapter 4 reveal that even consonantal phonemes, which, except the final nasal (/N/, hatsuon), have no corresponding orthographic signs in their native language writing system, play important roles in the perception of spoken Japanese by native listeners.

Further, orthographic influences on the phonological representations at the above two levels of processing were investigated.

It is concluded that phonological representations of the Japanese language are influenced by the orthography and writing system at the metaphonological level, at which Japanese native speakers paid the most attention to speech sounds. More specifically, findings show that the orthographic knowledge of the native language of Japanese (i.e., kana, syllabic phonograms) has a strong and pervasive influence.

However, no orthographic influence was observed at the preattentive level of speech processing, the level at which listeners pay the least attention to speech sounds. Here, the involvement of consonantal phonemes showed no influence of orthography or the writing system, namely kana. This result was different from the results of similar studies that used phoneme-based writing systems such as alphabets.

These conclusions from the studies described in Chapter 3 and Chapter 4 further lead to a general conclusion in relation to the levels of processing.

As was stated in Chapter 1 (1.4.2), in the levels-of-processing approach, the lowest level (i.e., the preattentive level of speech processing) functions for “perceptual” processing, including modular mandatory operations. This preattentive level is influenced only by early linguistic experience (before literacy acquisition), and the phonological representation of such a low level is not accessible directly to consciousness. In contrast, the highest level involves “conscious access” to the properties represented in the speech percept, and an “intentional analysis process” (hence meta-phonological activity or phonological awareness), and includes the processes that can be influenced by the knowledge related to literacy and attentional strategies.

The research work presented in Chapter 3 examined phonological representations of the Japanese language at the highest level. The findings show that the influence of orthographic knowledge, the knowledge related to literacy (i.e., that of kana), is consistent with the characteristics of the highest level.

Chapter 4 presented the research work examining the phonological representations at the lowest level, which is influenced only by early linguistic experiences before literacy acquisition. As stated earlier in this section, no orthographic influence was observed at this level, which means knowledge related to literacy did not have any influence. This finding is consistent with the notion of the lowest level.

Therefore, the general conclusion is that two of the three levels of the levels-of-processing approach appear to be supported by the two core research works in my dissertation on phonological representations of the Japanese language.

5.2 Implications for Further studies

There remains, however, one more level, the second level of “post-perceptual processes”, whose function is to re-elaborate the output of the perceptual module to take into account further relevant knowledge about the same objects or events, and which contributes to recognition. Further studies on phonological representations of the Japanese language are necessary at that second level to provide further support for the levels-of-processing approach.

5.2.1 Implications from the Research Work in Chapter 3

Since the first study in Chapter 3 was run with Japanese-like pseudowords, further investigation will have to assess whether such a behavior will also be observed with a non-native material, for instance with English-like pseudowords. However, such investigation should take into

account the fact that non-native listeners' performances might be biased by the way they perceive and/or conceptualize the stimuli, which itself depends on the phonological characteristics of their native language. For example, when the native language does not include consonant clusters in syllables, as it is the case in Japanese and in several Chinese languages, European words are typically broken up to that each consonant has its own syllable (Wang, 1973). This is quite evident for printed words. For example, Chinese renders the name "Clinton" with three characters representing three syllables: kuh-lin-ton (cf. Huang & Hanley, 1995). Epenthetic vowels are also frequent in Japanese written representation of European words. For example, the name "Clinton" is written with five characters:

Ku-ri-n-to-n (クリントン)

The same phenomenon of vowel epenthesis is observed as far as spoken language is concerned (for examples, see Shibatani, 1990). It seems to have profound implications in the domain of metaphonology and perhaps even in the perceptual domain (Dupoux, Kakehi, Hirose, Pallier, & Mehler, 1999). For example, as regards metaphonology, it has been shown that the ability to parse consonant clusters differs between Chinese and British children: whereas the latter perform initial phoneme deletion better on CVCC words, Hong Kong children perform better on CCVC words, most probably because these children implicitly introduce a vowel after the initial consonant in English words of this kind, thus transforming the phoneme into a syllable deletion task (Huan & Hanley, 1995).

Since one does not know whether such adaptations will hold true for the kind of population being tested, namely for highly proficient bilinguals and biliterates, any cross-language testing should use severely controlled material. In any case, the important point is that even with a native language-like material as the one used in the present study, the language-specific procedures adopted by the participants are powerful enough to allow a quite reasonable level of performance in metaphonemic analysis.

However, since bilingualism and biliteracy were confounded in the present study, further investigations should examine the possible effects of bilingualism *per se* on phonological awareness. As a matter of fact, several studies have suggested that second-language (L2 hereafter) learning can accelerate phonological awareness, but this notion has received little and even mixed empirical support, at least as far as phonemic analysis is concerned (e.g., Bruck and Genesee, 1995; Campbell & Sais, 1995; Rubin & Turner, 1989) rather than more general linguistic awareness (e.g., Bialystok, 1988, 1997; Galambos & Goldin-Meadow, 1990). Nevertheless, replication of present study with a group of fluent speakers of Japanese whose first orthography was not kana but alphabet (e.g., returnees from English-speaking countries, Japanese immigrants in Portuguese or Spanish-speaking countries in South America, such as Brazil or Peru, and so on and/or temporary workers visiting Japan from there) may help us disentangle the effects of bilingualism and those of biliteracy. Whereas the language-specific procedures used by the participants were based on their kana reading competence, it is at least possible that their bilingualism makes them particularly proficient in devising smart, alternative ways to achieve metaphonemic analysis.

Then, what will be found about Japanese adults' phonemic awareness in the acquisition of alphabetic spelling in a foreign language such as English or French? Since Japanese adults prefer to resort to their native language spelling resources, one can predict that a considerable number of poor readers of alphabetic spelling may exist among Japanese adults who are good at reading the moraic kana writing system.

Bassetti (2006) proposed that first language orthography influences the mental representations of L2 phonology in L2 learners who are beginners. Bassetti studied the phonological representations of Chinese rhymes in beginning learners of Chinese as a Foreign Language, using a phoneme counting task and a phoneme segmentation task. The results showed that learners did not count or segment the main vowel in those syllables where it is not represented in pinyin (romanization) orthographic representations.

Further psycholinguistic experiments on the performance level of

phonemic awareness should be conducted by using the same kinds of methodology that have been employed in research on alphabetic languages (e.g., rhymes and alliterations, phoneme deletion: Bryant, MacLean, Bradley, & Crossland, 1990; phoneme identification: Byrne & Fielding-Barnsley, 1990; phoneme counting: Castles, Holmes, Neath, & Kinoshita, 2003).

In relation to L2 learning and teaching, another interesting question is how a learner of the Japanese language who is a speaker of an alphabetic language would perform the same tasks if the instruction is given in Japanese by a Japanese native speaker. Furthermore, for cross-linguistic comparison, the same material and methodology should be applied to speakers of an alphabetic language such as English or French. It would also be interesting to apply the methodology to the considerable population of Japanese immigrants in South America, as well as the many temporary workers and their families living in Japan whose first language (L1 hereafter) is Portuguese or Spanish.

The relation between the acquisition of alphabetic literacy (or *romaji* in Japanese) and phonemic awareness of Japanese is still open to question. It will be necessary to compare with another Japanese adult group who have not acquired alphabetic literacy (or, if it is difficult to recruit such participants, those who are not familiar with reading alphabetic orthography) as a control group, as in de Gelder et al.'s (1993) experiment, which is discussed in Chapter 1 of this dissertation (p.35).

In fact, the method used in Experiment 3 was used in another study (not described here) involving 25 children aged 5 to 12 in Japan.¹ The vast majority of the children's answers corresponded to mora or syllable-based reversals, which were also found in the adults. The first-, second-, and third-graders tended not to understand how to respond in the task, and those up to grade four displayed very poor performance on phonemic reversals. Interestingly, however, the fifth- and the sixth-graders showed relatively good performance, and most of them, like the adult participants, employed kana and table-based strategies.

Further investigation is required concerning Japanese elementary school children, as well as preschoolers and even beginning learners of English as a foreign language in junior high school.

5.2.2 Implications from the Research Work in Chapter 4

As for the study in Chapter 4, contrary to present data and to the former results reported by Cutler and Otake (2002), which show that consonantal phonemes intervene in Japanese speech perception, Verdonschot, et al. (2011, see the end of 4.6, General Discussion) reported no priming effect for single segments in Japanese speech production. Whether this inconsistency reflects task differences (word reading vs. spoken word recognition) or a more fundamental difference between the functional units used in speech perception and speech production in Japanese should be investigated in future studies.

The present study showed that in Japanese speech perception, several sublexical units are functionally relevant at an early, preattentive stage of processing that seems unaffected by orthographic representations. Support for these conclusions may come from future studies comparing Japanese children of various levels of literacy with auditory target presentation (as in Experiment 2 in Chapter 4 on French speakers). In particular, it would be interesting to compare three groups of children: those who are not yet able to read either kana or kanji, those who read only kana, and those who already read both kana and kanji. If these three groups were to obtain similar results, and if these were similar to the adult results reported in the present work, the conclusions will be strongly reinforced.

Concerning L2 listening, the adult listening system of L1 is maladapted to differently structured input from their L2. L2 listening is likely to suffer from extensive unnecessary lexical activation and hence added competition. The mapping between the two vocabularies also adds difficulty, and the L1 vocabulary can be activated by L2 speech, too, which makes listening in a L2 noticeably harder than listening in the first (Cutler, 2012, p.303). Strange and Shafer (2008) noted the initial and continuing perceptual

difficulties of L2 learners and suggested a need for designing better instructional materials and tasks to improve perception of difficult L2 phonetic structures, as well as a need for cross-language research on the cognitive mechanisms involved in both L1 and L2 phonetic perception. They maintained that future studies using both brain and behavioral indices of perception, as well as laboratory training studies, should be designed to investigate the underlying mechanisms. This can also be applied to the L2 speech segmentation issue, which is implied from the research work outlined in Chapter 4.

As discussed in Chapter 1 (1.2.2), orthographic representations modulate the pronunciation of novel words in word learning situations. This has implications for adults acquiring a foreign language. In fact, it has been shown that L2 learners automatically generate orthographic forms for novel auditory words, based on the phonotactics of their native language (Johnston, McKague, & Pratt, 2004). When these forms are accurate, they can accelerate learning. During L2 learning, though, they might be inaccurate and impair learning.

Further, Detey and Nespoulous's (2008) study on Japanese learners of French suggested a possible influence of orthography on L2 syllabic representations, as the audiovisual and visual conditions triggered more epentheses than the auditory condition. Detey and Nespoulous concluded that the orthographic factor should not be neglected in L2 speech perception studies, loanword phonology, and interphonology research.

However, there is also a question of whether orthographic representations help or hurt L2 learning. Even training L2 learners with orthographic forms can have positive or negative effects on L2 learning, depending on the (in)congruence between grapheme-phoneme correspondences across the listeners' languages (i.e., Spanish as L1 and Dutch as L2; Escudero, Simon, & Mulak, 2014). By examining the effect of orthography on non-native word learning, Escudero et al. predicted that congruence between grapheme-phoneme correspondences across the listeners' languages could aid word recognition, while incongruence could hinder it. Although the study was concerned with the congruence of two

alphabetic orthographies, the results may provide a possibility for future studies of Japanese L2 learners to explore whether training L2 orthography will have positive or negative effects on L2 learning.

Further studies are needed to consider L2 listening acquisition in relation to native language (L1) speech perception and the influence of orthography.

We can predict that the present results for identifying the segmentation units, as well as the role of phoneme in the Japanese language at the level of preattentive speech perception, will help develop more natural and effective methods of L2 learning and teaching, as will be stated in the next section.

5.3 Future Directions

This research can be applied in the establishment of a new teaching method to support L2 literacy acquisition, that is, learning the orthography of English as a foreign language (EFL) or the alphabetic orthography of other languages.

In fact, phonological awareness of L2 facilitates L2 orthography learning. The findings of Sipra's (2003) study of Urdu learners of English as their L2 showed the significance of phonological awareness of English spellings to avoid wrong L2 phonology. Sipra recommended that phonetics and phonology should be part of the syllabus right from the elementary level, so that beginner level learners will be able to acquire some kind of phonological awareness before they start reading and writing (p.123). Even in Japan, there are a few research works on phonological awareness in L2 acquisition (experimental research works: Yuzawa, Mi., 2009; Yuzawa, Ma., 2007; error analysis: Stephan, 2013; proposals for English teaching: Jannuzi, 1998; Reynolds, 1998).

However, the application of phonological awareness training of L2 for Japanese literates has not been well developed (e.g., Goetry et al., 2005; Yuzawa & Yuzawa, 2013). Cheung (1999) examined the effects of phonological skill training and word reading performance in a biliterate

population whose L1 was Chinese (with a logographic non-alphabetic writing system) and whose L2 was English (with a phonographic alphabetic system) as L2. The findings of the study suggest that specialized segmentation training could contribute to promoting consonantal phonemic analysis, and that improved phonological skills could lead to better word reading in the later learned writing system. How to bridge between the spontaneous awareness of mora and the awareness of L2 syllable (e.g., English syllable) in Japanese literates remains unclarified.

There is an urgent need for studies that will provide insights for teachers and others who are involved in helping mora-oriented Japanese literates to acquire L2 phonological awareness, namely, that of phoneme and syllable in an alphabetic language.

Derwing and Munro (2005) observed that a growing body of literature on L2 speech has been published in journals that focus on speech production and perception (the research reported in Chapter 4 is published in one of those journals).² However, Derwing and Munro pointed out that such works are rarely cited or interpreted in teacher-oriented publications, and that the consequences of ignoring this important body of work are serious. They also noted that recent software for computer-aided instruction (hereafter CAI) appears to exploit multimedia capabilities rather than present linguistically and pedagogically good content.

Therefore, the studies described in my dissertation will shed light on L2 teaching and training methods and/or CAI applications for Japanese native speakers.³

Another possible application of the research work in Chapter 3 is in contributing to the development of a method of intervention and/or remediation for Japanese dyslexics⁴ and persons with reading disabilities.

Recent research works have been challenging the widespread belief that dyslexia, especially phonological dyslexia, is largely absent among Japanese children.⁵ Yamada and Banks (1994) found that 6% of their sample of 125 fourth-graders qualified as dyslexics and that dyslexia exists to a significant degree in Japan. Oishi and Saito (1999) found that phonological

awareness (word reversals and mora-counting) of seven children who were developmental dyslexics was far below that of the control group. Oishi and Saito compared the results of two kinds of word retrieval tasks, and found a discrepancy in the number of words retrieved in the two tasks. That is, the number of words with a mora-sound retrieved was inferior to that of words with a categorical meaning, and such a discrepancy was not observed among the controls. These results indicated that the seven dyslexics had a developmental delay in phonology, revealing that development of phonology plays an important role in learning hiragana as well. More recently, Tanaka-Welty (2005) found that five boys with language-based learning disabilities (LLD) scored significantly worse in word identification, kana spelling, syllable deletion, and kana matrix completion than the normally developing (ND) children. Her results indicated that the problems of the boys with LLD appeared even in kana decoding, and that their disabilities might be based on phonological processing deficits. Findings of Tanaka et al. (2006) further support the above conclusion, as the problems of children with reading disabilities were related to kana decoding and their disabilities might be based on phonological processing deficits. Similarly, Seki et al. (2008) examined the abilities of Japanese dyslexic children on different types of phonological tasks in which kana (the transparent, mora-based phonogram) was used. They found that Japanese dyslexics showed deficits of phonological awareness at both the mora and the phoneme levels, and concluded that phonological awareness must be crucial for acquiring the ability of decoding phonograms, including Japanese kana.

Thus, it has been proved that dyslexics and persons with reading disabilities in phonological processing actually exist in Japan. It is implied that the present research, especially the study on phonological awareness among Japanese (Chapter 3), would help in the remediation and even screening of dyslexics and persons with reading disabilities in Japan. For example, the reversal task used in Chapter 3 has already been applied to one of the test batteries used in a study on acquired alexia (Sasanuma, 1996).

The study described in Chapter 3 will also contribute to developing a new system of CAI for purposes of screening and remediation of dyslexics and persons with reading disabilities. A considerable number of studies have shown the positive effects of computer applications for training phonological awareness in the English language as L1, and in some other languages.⁶

For instance, Barrons et al. (1992) trained phonological awareness skills of English as L1 in normally developed preliterate children by using a DECtalk speech synthesizer. Barrons et al. found that a combination of high letter-sound knowledge and print feedback facilitates awareness of phonemes among children who cannot yet read or spell.⁷ Wild (2009) investigated the use of CAI for practicing phonological awareness skills with beginning readers. The statistical analysis showed a significant learning advantage of children in the computer-based group compared to the paper-based control group, although this learning advantage was less strong for the orthographic application of their phonological skills. Lonigan et al. (2003) evaluated the use of CAI to provide preschoolers at risk for reading problems with training in phonological sensitivity skills. They found that those exposed to CAI made significantly greater gains in rhyming and elision skills compared to the control group. Lonigan et al. concluded that their results provided preliminary support for CAI to improve the phonological sensitivity skills of preschoolers who are at risk. Jiménez et al. (2007) assessed the effects of four reading-training procedures for Spanish children with reading disabilities in a transparent orthography of Spanish as their L1, with the aim of examining the effects of different spelling-to-sound units in computer speech-based reading. In the areas of word recognition, reading comprehension, phonological awareness, and orthographic and phonological tasks, the results suggest the importance of training phonological processes in improving word decoding in children with dyslexia who learn in a consistent orthography.

As for Japanese dyslexics and persons with reading difficulties, I was not able to find any computer applications for their remediation, especially CAI for the acquisition of Japanese phonology-orthography correspondence.

Even if they exist, the number could be very small. The findings from the present research will help develop CAI instruments for the remediation of Japanese dyslexics and persons with reading difficulties.

Some research on text comprehension in a second language (Morishima, 2013) indicates that there are limitations on the cognitive resources allocated for discourse-level processes in L2 comprehension. The lower-level language processes, such as lexical access, involved in the construction of surface form and text base are more cognitively demanding in L2 comprehension than in L1 comprehension. If the research in my dissertation is applied to ways of improving and/or training segmentation skills and phonological awareness in L2, we can expect a decrease in cognitive loads for lexical access and amount of time spent at the level of word recognition. With speech segmentation training and phonological awareness enhancement, Japanese L2 learners will be able to allocate more cognitive resources to the higher levels such as text comprehension. Thus, the studies make a significant contribution not only to our understanding of phonology and orthography, but also to research concerning the higher discourse level of L2 acquisition.

To sum up, I believe that the studies presented in this dissertation will make important contributions to helping not only Japanese native speakers who use the mora-based kana writing system to learn foreign languages with phoneme-based alphabetic writing systems (such as English), but also Japanese dyslexics and persons with reading disabilities who have difficulties in kana acquisition that are related to their impaired phonological processing.

BIBLIOGRAPHY

- Adams, M. J. (1990). *Beginning to read: thinking and learning about print*. Cambridge, MA: MIT Press.
- Adrián, J. A., Alegria, J., & Morais, J. (1995). Metaphonological abilities of Spanish illiterate adults. *International Journal of Psychology*, 30, 329-353.
- Alegria, J., Pignot, J. & Morais, J. (1982). Phonetic analysis of speech and memory codes in beginning readers. *Memory and Cognition*, 10, 451-456.
- Allport, A. (2011). Attention and integration, In C. Mole, D. Smithies, & W. Wu (Eds.) *Attention: philosophical and psychological essays* (pp.24-59), Oxford: Oxford University Press.
- Amano, K. (1970). Go no on-in kozo no bunseki-koi no keisei to kanamoji no yomi no gakushu [Formation of the act of analyzing the phonemic structure of words and its relation to learning Japanese syllabic characters]. *Japanese Journal of Educational Psychology*, 18, 76-89.
- Amano, K. (1986). *Kodomo no kanamoji no shutoku-katei [The acquisition process of Japanese syllabic characters in children]*. Tokyo: Akiyama-shoten.
- Amano, K. (1988). On-in bunseki to kodomo no literacy no shutoku [Acquisition of phonemic analysis and literacy of children]. *Annual Report of Educational Psychology in Japan*, 27, 142-164.
- Backman, J. (1983). The role of psycholinguistic skills in reading acquisition: A look at early readers. *Reading Research Quarterly*, 28, 466-479.
- Barrons, R. W., Golden, J. O., Seldon, D. M., Tait, C. F., Marmurek, H. H. C., & Haines, L. P. et al. (1992). Teaching prereading skills with a talking computer: Letter-sound knowledge and print feedback facilitate nonreaders' phonological awareness training. *Reading and Writing: An Interdisciplinary Journal*, 4, 179-204.
- Bassetti, B. (2006). Orthographic input and phonological representations in learners of Chinese as a Foreign Language, *Written Language and*

Literacy, 9(1), 95-114.

- Bertelson, P. (1987). The onset of literacy: Liminal remarks. In P. Bertelson (Ed.) *The onset of literacy: Cognitive process in reading acquisition* (pp.1-30). Cambridge (MA): MIT Press.
- Berthoud-Papandropoulou, I. (1978). An experimental study of children's ideas about language. In A. Sinclair, R. J. Jarvella & W. J. M. Levelt (Eds.), *The Child's Conception of Language* (pp. 55-65). New York: Springer-Verlag.
- Bialystok, E. (1985) Toward a Definition of Metalinguistic Skill, *Merril Palmer Quarterly*, 31(3) 229-251.
- Bialystok, E. (1986) Factors in the Growth of Linguistic Awareness, *Child Development*, 57, 498-510.
- Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Developmental Psychology*, 24, 560-567.
- Bialystok, E. (1997). Effects of bilingualism and biliteracy on children's emerging concepts of print. *Developmental Psychology*, 33, 429-440.
- Birdsong, D. (1989). *Metalinguistic performance and interlinguistic competence*. Berlin ; New York : Springer-Verlag.
- Boersma, P. (2001). Praat: a system for doing phonetics by computer. *Glott International*, 5, 341-345.
- Boersma, P. & Hamann, S. (2009). Introduction: models of phonology in perception. In Boersma & Hamann (Eds.), *Phonology in Perception*, The Hague: Mouton de Gruyter.
- Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer [Computer program]. Version 5.3.42, retrieved 2 March 2013 from <http://www.praat.org/>.
- Bruck, M. & Genesee, F. (1995). Phonological awareness in young second language learners. *Child Language*, 22, 307-324.
- Bryant, P. E., MacLean, M., Bradley, L. L., & Crossland, J. (1990). Rhyme and alliteration, phoneme deletion, and learning to read. *Developmental Psychology* 26, 429-438.
- Bürki, A., Spinelli, B., & Gaskell, M. G. (2012). A written word is worth a thousand spoken words: the influence of spelling on spoken-word

- production. *Journal of Memory and Language*, 67, 449-467.
- Byrne, B. and Fielding-Barnsley, R. (1990). Acquiring the alphabetic principle: A case for teaching recognition of phoneme identity. *Journal of Educational Psychology*, 82, 805-812.
- Campbell, R.; Butterworth, B. (1985). Phonological dyslexia and dysgraphia in a highly literate subject: A developmental case with associated deficits of phonemic processing and awareness. *The Quarterly journal of experimental psychology. A, Human experimental psychology* 37 (3), 435-475.
- Campbell, R., & Sais, E. (1995). Accelerated metalinguistic (phonological) awareness in bilingual children. *British Journal of Developmental Psychology*, 13, 61-68.
- Carrillo, M. (1994). Development of phonological awareness and reading acquisition: A study in Spanish language. *Reading and Writing*, 6, 279-298.
- Castles, A., Holmes, V. M., Neath, J. & Kinoshita, S. (2003). How does orthographic knowledge influence performance on phonological awareness tasks? *Quarterly Journal of Experimental Psychology*, 56A, 445-467.
- Castro, S. -L., Vicente, S., Morais, J., Kolinsky, R., & Cluytens, M. (1995). Segmental representation of Portuguese in 5- and 6- year olds: Evidence from dichotic listening. In I. Hub Faria & J. Freitas (Eds.), *Studies on the acquisition of Portuguese* (pp. 1-16). Lisbon: Colibri.
- Cheung, H. (1999). Improving phonological awareness and word reading in a later learned alphabetic script. *Cognition*, 70, 1-26.
- Church, K.W. (1987). Phonological parsing and lexical retrieval. *Cognition*, 25, 53-69.
- Clark, E. V. (1978). Awareness of language: Some evidence from what children say and do. In A. Sinclair, R. J. Jarvella, & W. J. M. Levelt (Eds.), *The Child's conception of language*. New York, NY: Springer-Verlag.
- Coleman, J. (1998). *Phonological Representations: Their names, forms and powers*. Cambridge: Cambridge University Press.

- Content, A. (1984). L'analyse phonétique explicite de la parole et l'acquisition de la lecture. *L'Année Psychologique*, 84, 555-572.
- Content, A., Kolinsky, R., Morais, J., & Bertelson, P. (1986). Phonetic segmentation in prereaders: effect of corrective information. *Journal of Experimental Child Psychology*, 42, 49-72.
- Content, A., Meunier, C., Kearns, R. K., & Frauenfelder, U. H. (2001). Sequence detection in pseudowords in French: Where is the syllable effect? *Language and Cognitive Processes*, 16, 609-636.
- Crystal, D. (1991). A dictionary of linguistics and phonetics. Oxford: Basil Brackwell.
- Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*, Cambridge (MA): MIT Press.
- Cutler, A., Mehler, J., Norris, D., & Seguí, J. (1983). A language specific comprehension strategy, *Nature*, 304, 159-160.
- Cutler, A., & Otake, T. (1994). Mora or phoneme? Further evidence for language-specific listening. *Journal of Memory and Language*, 33, 824-844.
- Cutler, A., & Otake, T. (2002). Rhythmic categories in spoken-word recognition. *Journal of Memory and Language*, 46, 296-322.
- Cutting, J. E. (1976). Auditory and linguistic processes in speech perception: Inference from six fusions in dichotic listening. *Psychological Review*, 83, 114-140.
- Day, R. S. (1968). Fusion in dichotic listening. Dissertation Abstracts International, 29, 2649B. (University Microfilms No. 69-211)
- de Gelder, B., Vroomen, J., & Bertelson, P. (1993). The effects of alphabetic-reading competence on language representation in bilingual Chinese subjects. *Psychological Research*, 55, 315-321.
- Dell, G. S., & Reich, P. A. (1980). Toward an unified theory of slips of the tongue. In V.A. Fromkin (Ed.), *Errors in Linguistic Performance: Slips of the Tongue, Ear, Pen, and Hand* (pp. 273-286). New York: Academic Press.

- Denny, J. P. (1991). Rational thought in oral culture and literate decontextualization. In D. R. Olson & N. Torrance (Eds), *Literacy and Orality* (pp. 66-89). New York, NY: Cambridge University Press.
- Derwing, T. M. & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 39(3), 379-397.
- Detey, S & Nespoulous, J-L. (2008). Can orthography influence second language syllabic segmentation?: Japanese epenthetic vowels and French consonantal clusters, *Lingua*, 118, 66-81.
- Dupoux, E, Kakehi, K., Hirose, H., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1568-1578.
- Dupoux, E., & Mehler, J. (1992). Unifying awareness and on-line studies of speech: A tentative framework. In J. Alegria, D. Holender, J. Junça de Moraes, & M. Radeau (Eds.), *Analytic Approaches to Human Cognition* (pp. 59-75). Amsterdam: North-Holland.
- Ehri, L. C. (1984). How orthography alters spoken language competencies in children learning to read and spell. In J. Downing & R. Valtin (Eds.), *Language awareness and leaning to read* (pp. 119-147). New York: Springer-Verlag.
- Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, 36, 250-287.
- Ehri, L. C., & Wilce, L. S. (1979). The mnemonic value of orthography among beginning readers. *Journal of Educational Psychology*, 71, 26-40.
- Ehri, L. C., & Wilce, L. S. (1980). The influence of orthography on reader's conceptualization of the phonemic structure of words. *Applied Psycholinguistics*, 1, 371-385.
- Ehri, L. C., & Wilce, L. S. (1986). The influence of spelling on speech: are alveolar flaps /d/ or /t/? In D. Yaden & S. Templeton (Eds.), *Metalinguistic awareness and beginning literacy* (pp. 101-114).

- Portsmouth: Heinemann.
- Elkonin, D. B. (1973). U.S.S.R. In J. Downing (ed.). *Comparative Reading*. New York: MacMillan.
- Ellis, A. W. (1993) *Reading, writing and dyslexia : a cognitive analysis*. Hove : Lawrence Erlbaum.
- Endo, M. (1991). Nihonjin yoji no in no kanjusei to yoon hyokiho no shutoku [Phonological awareness of Japanese young children and learning to read and write yoo-on]. *Japanese Journal of Educational Psychology*, 39, 448-454.
- Escudero, P., Simon, E., & Mulak, K. E. (2014). Learning words in a new language: orthography doesn't always help. *Bilingualism: Language and Cognition*, 17, 384-395.
- Feldman, C. F. (1991). Oral metalanguage, In D. R. Olson & N. Torrance (Eds), *Literacy and Orality* (pp. 47-65). New York, NY: Cambridge University Press.
- Fernandes, T., Kolinsky, R., Ventura, P. (2010). The impact of attention load on the use of statistical information and coarticulation as speech segmentation cues. *Attention, Perception & Psychophysics*, 72, 1522-1532.
- Fernandes, T., Ventura, P., & Kolinsky, R. (2007). Statistical information and coarticulation as cues to word boundaries: A matter of quality of the signal. *Perception and Psychophysics*, 69, 856-864.
- Fodor, J. A. (1972). Some reflections on L. S. Vygotsky's Thought and Language. *Cognition*, 1, 83-95.
- Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.
- Frazier, L. (1987). Structure in auditory word recognition. *Cognition*, 25, 157-187.
- Frauenfelder, U. H., Seguí, J., & Dijkstra, T. (1990). Lexical effects in phonemic processing: Facilitatory or inhibitory? *Journal of Experimental Psychology: Human Perception and Performance*, 16, 77-91.
- Fromkin, V. A. (1973). (Ed.), *Speech Errors as Linguistic Evidence*. The

- Hague: Mouton.
- Fujimura, O. (1966). Kotoba no kagaku [Science of language]. *Shizen*, 21, 48-53.
- Fujimura, O. (2007). *Onseikagaku genron: Gengo no honshitsu wo kangaeru* [Theory of speech science: Consideration on the essence of language], Tokyo: Iwanami Shoten.
- Fukumori, T. (2010). *Kisokara no nihongo onseigaku* [Basic study of Japanese phonetics]. Tokyo: Tokyodoshuppan.
- Galambos, S. J., & Goldin-Meadow, S. (1990). The effects of learning two languages on levels of metalinguistic awareness. *Cognition*, 34, 1-56.
- Gathercole, S. E. & Baddeley, A. D. (1993). *Working memory and language*. Hove: Lawrence Erlbaum.
- Goetry, V., Urbain, S., Morais, J., & Kolinsky, R. (2005). Paths to phonemic awareness in Japanese: Evidence from a training study, *Applied Psycholinguistics*, 26(2), 285-309
- Goldinger, S. D. (1996). Words and voices: episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22 (5), 1166-1183.
- Goldsmith, J. A. (1993). Harmonic phonology in J. A. Goldsmith (Ed.) *The last phonological rule: reflections on constraints and derivations*. Chicago: The University of Chicago Press.
- Grabe, E., & Lowe, E.L. (2002). Durational variability in speech and the rhythm class hypothesis In: Gussenhoven, C., Warner, N. (Eds.), *Papers in Laboratory Phonology 7*, (pp. 515-546). Berlin: Mouton de Gruyter.
- Hakes, D. (1980). *The development of metalinguistic abilities in children*. Berlin: Springer-Verlag.
- Haraguchi, S. (1999). Accent, In N. Tsujimura (Ed.), *The Handbook of Japanese Linguistics* (pp.1-30). Oxford: Blackwell Publishers.
- Hardison, D. M., & Motohashi Saigo, M. (2010). Development of perception of second language Japanese geminates: Role of duration, sonority, and segmentation strategy. *Applied Psycholinguistics*, 31, 81-99.

- Hatano, G., & Inagaki, K. (1992). Yoji no on-in ishiki: sono 2. Shiritori wo toshite [Phonological awareness in Japanese children, No.2: Through the verbal Game, shiritori]4. *Proceedings of the 56th Annual Convention of the Japanese Psychological Association*. Doshisha University, Kyoto, 813.
- Hattori, S. (1990). *On-inron to seishoho [Japanese phonology and orthography. Rev. Ed.]*. Tokyo, Japan: Taishukan-shoten.
- Hayes, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry*, 20, 253-306.
- Holm, A., & Dodd, B. (1996). The effect of first written language on the acquisition of English literacy. *Cognition*, 59, 119-147.
- Huang, H. S., & Hanley, J. R. (1995). Phonological awareness and visual skills in learning to read Chinese and English. *Cognition*, 54, 73-98.
- Hutson, B. A. (1979). Macro-theory and micro-theory in the study of concepts about language and reading. *Paper presented at the University of Victoria/IRA Research Seminar on Linguistic Awareness and Learning to Read*, Victoria, B. C..
- Hyman, L. (1985). *A theory of phonological weight*. Dordrecht: Foris.
- Imada, S. (1990). Hatsuon [Pronunciation]. In the Japan Foundation Japanese Language Institute (Ed.), *Kyoshiyo nihongo handbook series*, vol. 6. Tokyo: Bonjin-sha.
- Inaba, S. (1998). Morae, syllables, and feet in Japanese. *Proceedings of Language, Information and Computation (PACLIC12)*, National University of Singapore, 106-117.
- Inagaki, K., & Hatano, G. (1992). Yoji no on-in ishiki: sono1. Go no bunkai wo toshite [Phonological awareness in Japanese children, No.1. Through a segmentation task]. *Proceedings of the 56th Annual convention of the Japanese Psychological Association*. Doshisha University, Kyoto, 812.
- 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.
- Ishizaki, S. (2013). Human Speech Analysis Process, *Unpublished lecture handout of Pattern Information Processing 10*, Keio University.

- Itô, J. (1990). Prosodic minimality in Japanese. *Chicago Linguistic Society 26. II: Papers from the parasession on the syllable in phonetics and phonology* (pp. 213-239). Chicago: CLS.
- Jannuzi, C. (1998). Key Concepts in FL Literacy: Phonemic Awareness. *Literacy Across Cultures, February 1/2*,
<http://www2.aasa.ac.jp/~dcdycus/LAC98/FEB98/PHONAWA1.HTM>
- Jiménez, J. E., Hernández-Valle, I., Ramírez, G., Ortiz, M. R., Rodrigo, M., Estévez, A., O'Shanahan, I., Garcia, E., & Trabaue, M. L. (2007). Computer speech-based remediation for reading disabilities: The size of spelling-to-sound unit in a transparent orthography. *The Spanish Journal of Psychology, 10*(1), 52-67.
- Johns, J. L. (1984). Students' perceptions of reading: Insights from research and pedagogical implications. In J. Downing & R. Valtin (Eds.), *Language awareness and leaning to read* (pp. 57-77). New York: Springer-Verlag.
- Johnston, M., McKague, M., & Pratt, C. (2004). Evidence for an automatic orthographic code in the processing of visually novel word forms. *Language and Cognitive Processes, 19*, 273-317.
- Jones, D. (1993). Objects, streams, and threads of auditory attention. In A. Baddeley & L. Weiskrantz (Eds.), *Attention: selection, awareness, and control: A tribute to Donald Broadbent* (pp.87-104). Oxford: Oxford University Press.
- Kamio, A., & Tonoike, S. (1979). Iimachigai no gengogaku [Linguistics of speech errors]. In K. Imai (Ed.), *Gengo-shogai to gengo-riron [Speech disorders and linguistic theories]*(pp.272-308), Tokyo, Japan: Taishukan-shoten.
- Kashino, M. (1991). Perceptual integration of distributed phonemic cues. *Progress Report of April 1990 through March 1991*. Information Science Research Lab. Tokyo: NTT Basic Research Labs. Kashino, M. (1991). Perceptual integration of distributed phonemic cues. *Progress Report of April 1990 through March 1991*. Information Science Research Lab. Tokyo: NTT Basic Research Labs.
- Kasuya, H., & Sato, S. (1990). CHOON, niren-boin to niyu-boin: Nihongo

- washa to eigo washa no baai [Long vowels, double vowels, and diphthongs: In the case of Japanese native speakers and English native speakers]. In M. Sugito (Ed.), *Koza nihongo to nihongo kyoiku, dai 3-kan: Nihongo no onsei, on-in (ge) [the Japanese language and its education, Vol.3, Speech sounds and phonological issues of the Japanese language (No. 2)]* (pp.178-197). Tokyo: Meiji-shoin.
- Katada, F. (1990). On the representation of morae: Evidence from a language game. *Linguistic Inquiry*, 21, 641-646.
- Kawahara, S. (2007). Half rhymes in Japanese rap lyrics and knowledge of similarity. *Journal of East Asian Linguistics*, 16, 113-144.
- Kawahara, S. (in press). The phonetics of obstruent geminates, sokuon. In H. Kubozono (Ed.), *The Mouton Handbook of Japanese Language and Linguistics*. Berlin: Mouton de Gruyter.
- Kawahara, S. & Shinohara, K. (2009a). Calculating vocalic similarity through puns. *Onsei Kenkyuu [Journal of the Phonetic Society of Japan]*, 13, 101-110.
- Kawahara, S. & Shinohara, K. (2009b). The role of psychoacoustic similarity in Japanese puns: A corpus study. *Journal of Linguistics*, 45, 111-138.
- Kerstholt, M. T., Van Bon, W. H. J. & Schreuder, R. (1994). Training in phonemic segmentation: the effects of visual support. *Reading and Writing: An Interdisciplinary Journal*, 6, 361-385.
- Koizumi, T. (1978). *Nihongo no seishoho [Japanese Orthography]*. Tokyo: Taishukan-shoten.
- Koizumi, T. (1989). Onsei to on-in [Speech sounds and phones of the Japanese language]. In M. Sugito (Ed.), *Koza nihongo to nihongo kyoiku, dai 2-kan: Nihongo no onsei, on-in (jo) [the Japanese language and its education, vol. 2, Speech sounds and phonological issues of the Japanese language]* (pp. 1-20). Tokyo: Meiji-shoin.
- Kolinsky, R. (1998). Spoken word recognition: A stage-processing approach to language differences. *The European Journal of Cognitive Psychology*, 10, 1-40.
- Kolinsky, R. (2006). Niveaux de traitement des mots parlés[Levels of

processing of spoken words], *Dissertation préparée en vue de l'examen pour l'obtention du grade d'Agrégé de l'Enseignement Supérieur [Unpublished dissertation prepared for the purpose of the examination for obtaining the grade of qualified professor for teaching at the superior level in Belgian universities]*.

- Kolinsky, R. (in press). How learning to read influences language and cognition. In A. Pollatsek & R. Treiman (Eds.), *The Oxford Handbook of Reading*. New York: Oxford University Press.
- Kolinsky, R., Cary, L., & Morais, J. (1987). Awareness of words as phonological entities: the role of literacy. *Applied Psycholinguistics*, 8, 223-232.
- Kolinsky, R. , & Morais, J. (1996). Migration in speech recognition. *Language and Cognitive Processes (special issue, A Guide to Spoken Word Recognition Paradigms)*, 11, 611-619.
- Kolinsky, R. , Morais, J. , & Cluytens, M. (1995). Intermediate representations in spoken word recognition: Evidence from word illusions. *Journal of Memory and Language*, 34, 19-40.
- Komatsu, H. (1981). *Nihongo no on-in [Japanese phonological issues]*. Tokyo: Shuokoron-sha.
- Kondo, Y. & Arai, Y. (1998). Prosodic constraint on V-to-V coarticulation in Japanese: A case of bimoraic foot. *On-in Kenkyuu [Phonological Studies]*, 1, 75-82.
- Kondo, Y. (2006). Within-word prosodic constraint on coarticulation in Japanese. *Language and Speech*, 49(3), 393-416.
- Kubo, T. (2003). On-in kenkyuu no dookoo to tenboo 2; Gendaigo chuusshin [Current trends and perspectives in phonology; Modern Japanese] in Z. Uwano (Ed.), *Onsei, on-in [Phonetics, Phonology]* (pp.239-248). Tokyo : Asakura Publishing.
- Kubozono, H. (1989). The mora and syllable structure in Japanese: evidence from speech errors. *Language and Speech*, 32, 249-278.
- Kubozono, H. (1992). Nihongo no moraeono yakuwari to tokusei [Japanese mora: Its role and properties]. In S. Haraguchi (Ed.), *Nihongo no mora to onsetsu-kozo ni kansuru sogoteki kenkyu. [A comprehensive study on*

- the structure of the Japanese mora and syllable*]. (pp. 48-61). Nagoya: Nanzan University.
- Kubozono, H. (1995). *Gokeisei to On'in Koozoo [Word Formation and Phonological Structure]*. Tokyo: Kuroshio.
- Kubozono, H. (1998). On the universality of mora and syllable. *Journal of the Phonetic Society of Japan*, 2, 5-15.
- Kubozono, H. (1999a). Nihongo no onsei [Japanese speech sounds]. Tokyo: Iwanami Shoten.
- Kubozono, H. (1999b). Kayoo ni okeru mora to onsetsu [Mora and syllables in Japanese songs], In Onsei-Bunpo Kenkyukai [Spoken Language Working Group] Ed., *Bunpo to onsei [Speech and grammar] II*. (pp.241-260). Tokyo: Kuroshio.
- Kubozono, H. (1999c). Mora and syllable, In N. Tsujimura (Ed.), *The Handbook of Japanese Linguistics*. (pp.31-61). Oxford: Blackwell Publishers.
- Kubozono, H. (2003). The syllable as a unit of prosodic organization in Japanese. In C. Féry & R. van de Vijver (Eds.), *The syllable in optimality theory*, (pp. 99-122). Cambridge: Cambridge University Press.
- Kubozono, H. (2006). Where does loanword prosody come from?: A case study of Japanese loanword accent. *Lingua*, 116, 1140-1170.
- Kubozono, H., & Ota, S. (1998). *On'in koozoo to Akusento [Universality and specificity of phonological structure]*. Tokyo: Kenkyusha.
- Kurtz, B.E. (1991) Cognitive and Metacognitive Aspects of Text Processing, In G. Denhiere and J-P. Rossi (Eds.) *Text and Text Processing*. North-Holland, Amsterdam.
- Labrune, L. (2006). *La phonologie du japonais*. Leuven: Peeters.
- Labrune, L. (2012a). Questioning the universality of the syllable: evidence from Japanese, *Phonology*, 29, 113-152.
- Labrune, L. (2012b). *The phonology of Japanese*. Oxford: Oxford University Press.
- Ladefoged, P. (1975). *A course in phonetics*. New York: Harcourt Brace Jovanovich.

- Lennon, J. & McCartney, J. P. (1966). Lyrics of Eleanor Rigby. Sony/ATV Music Publishing LLC.
- Levelt, W. J. M., Sinclair, A. & Jarvella, R. J. (1978). Causes and functions of linguistic awareness in language acquisition: some introductory remarks. In A. Sinclair, R. J. Jarvella & W. J. M. Levelt (Eds.), *The Child's Conception of Language* (pp.1-14). New York: Springer-Verlag.
- Liberman, I. Y. (1971). Basic research in speech and the lateralization of language: Some implications for reading, *Bulletin of the Orton Society*, 21, 71-87.
- 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.
- Lonigan, C. J., Driscoll, K., Phillips, B. M., Cantor, B. G., Anthony, J. L., & Goldstein, H. (2003). A computer-assisted instruction phonological sensitivity program for preschool children at-risk for reading problems. *Journal of Early Intervention*, 25(6), 248-262.
- Lundberg, I. (1978). Aspects of linguistic awareness related to reading. In A. Sinclair, R. J. Jarvella, & W. J. M. Levelt (Eds.), *The child's conception of language* (pp. 83-96). Berlin Heidelberg New York, NY: Springer-Verlag.
- Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia, *Annals of Dyslexia*, 53, 1-14.
- Machida, Y. & Asano, K. (Ed.) (1962). *Warabe-uta [Nursery songs]*, Tokyo : Iwanami Shoten.
- McCawley, J. D. (1968). The phonological component of a grammar of Japanese. The Hague: Mouton.
- McCawley, J. (1977). Accent in Japanese. In L. Hyman (Ed.), *Studies in stress and accent*. (pp.261-302). Los Angeles: University of Southern California.
- MacKay, D. G. (1970). Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia*, 8, 323-350.
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide. (2nd edition)*. Mahwah, NJ: Erlbaum.

- Makita, K. (1968). The rarity of reading disability in Japanese children. *American Journal of Orthopsychiatry*, 38(4), 599-614.
- Mann, V. (1986). Phonological awareness: The role of reading experience. *Cognition*, 24, 65-92.
- Mann, V. (1991). Phonological awareness and early reading ability: one perspective. In D. J. Sawyer & B. J. Fox (Eds.), *Phonological awareness in reading: The evolution of current perspectives* (pp.191-215). New York: Springer-Verlag.
- Marcel, A. J. (1983). Conscious and unconscious perception: An approach to the relations between phenomenal experience and perceptual processes. *Cognitive Psychology*, 15, 238-300.
- Mattingly, I.G.(1972). Reading, the linguistic process and linguistic awareness, In J.F. Kavanagh, I.G. Mattingly (Eds.), *Language by ear and by eye: The relationship between speech and reading*, Cambridge, MA: MIT Press.
- Mattys, S. L. (2004). Stress versus coarticulation: Towards an integrated approach to explicit speech segmentation. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 397-408.
- Mattys, S. L., Brooks, J., & Cook, M. (2009). Recognizing speech under a processing load: Dissociating energetic from informational factors. *Cognitive Psychology*, 59, 203-243.
- Mattys, S. L., & Melhorn, J. F. (2005). How do Syllables Contribute to the Perception of Spoken English? Insight from the Migration Paradigm. *Language and Speech*, 48, 223-253.
- Mattys, S. L. & Samuel, A. G. (1997). How lexical stress affects speech segmentation and interactivity: Evidence from the migration paradigm. *Journal of Memory and Language*, 36, 87-116.
- Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology: General*, 134, 477-500.
- Mazuka, R. (2009) Acquisition of linguistic rhythm and prosodic bootstrapping hypothesis, *Journal of the Phonetic Society of Japan*, 13 (3), 19-32.

- McClelland, J. L., & Mozer, M. C. (1986). Perceptual interactions in two-word displays: familiarity and similarity effects. *Journal of Experimental Psychology: Human Perception and Performance*, 12, 18-35.
- McQueen, J. M., Otake, T., & Cutler, A. (2001). Rhythmic cues and possible-word constraints in Japanese speech segmentation. *Journal of Memory and Language*, 45, 103-132.
- Mehler, J. & Christophe, A. (1995) Maturation and Learning of Language in the First Year of Life. In M. S. Gazzaniga (Ed) *The Cognitive Neurosciences*. (pp. 943-953). Cambridge, MA.: MIT Press.
- Mehler, J., Dommergues, J. Y., Frauenfelder, U., & Seguí, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, 20, 298-305.
- Mester, A. R. (1990). Patterns of truncation. *Linguistic Inquiry*, 21(3), 478-485.
- Mole, C., Smithies, D., & Wu, W. (2011). Introduction, In C. Mole, D. Smithies, & W. Wu (Eds.) *Attention: philosophical and psychological essays* (pp.xi-xx), Oxford: Oxford University Press.
- Morais, J. (1987). Segmental analysis of speech and its relation to reading ability, *Annals of Dyslexia*, 37(1), 126-141.
- Morais, J. (1991). Phonological awareness: A bridge between language and literacy. In D. J. Sawyer & B. J. Fox (Eds.), *Phonological awareness in reading: The evolution of current perspectives*. (pp. 31-71). New York: Springer-Verlag.
- Morais, J. (1993). Phonemic awareness, language and literacy. In R. M. Joshi & C. K. Leong (Eds.), *Reading disabilities: Diagnosis and component processes* (pp. 175-184). Dordrecht: Kluwer.
- Morais, J., Alegria, J., & Content, A. (1987). The relationships between segmental analysis and alphabetic literacy: An interactive view. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, 7, 415-438.
- Morais, J., Bertelson, P., Cary, L., & Alegria, J. (1986). Literacy training and speech segmentation. *Cognition*, 24, 45-64.
- Morais, J., Cary, L., & Alegria, J. & Bertelson, P. (1979). Does awareness of

- speech as a sequence of phones arise spontaneously? *Cognition*, 7, 323-331.
- Morais, J., Castro, S.-L., Scliar-Cabral, L., Kolinsky, R., & Content, A. (1987). The effects of literacy on the recognition of dichotic words. *Quarterly Journal of Experimental Psychology*, 39 A, 451-465.
- Morais, J., Content, A., & Alegria, J. (1987). Segmental awareness: Respectable, useful, and almost always necessary. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, 7, 530-556.
- Morais, J., & Kolinsky, R. (1992). In search of a paradigm for the study of speech perceptual codes. *I Jornadas de Estudo dos Processos Cognitivos da Sociedade Portuguesa de Psicologia* (pp. 113-135). Lisbonne: Astoria.
- Morais, J., & Kolinsky, R. (1994; 1995). Perception and awareness in phonological processing: The case of the phoneme. *Cognition*, 50, 287-297. Also in J. Mehler & S. Franck (Eds.), *Cognition on cognition* (pp. 349-359). Cambridge: MIT Press.
- Morais, J., & Kolinsky, R. (2001). The literate mind and the universal human mind. In E. Dupoux (Ed.), *Language, brain and cognitive development: Essays in honor of Jacques Mehler* (pp.463-480). Cambridge, Mass: M.I.T. Press.
- Morais, J., Kolinsky, R., Alegria, J., & Scliar-Cabral, L. (1998). Alphabetic literacy and psychological structures. *Letras de Hoje*, 33, 61-79.
- Morais, J., Kolinsky, R., & Nakamura, M. (1996). The psychological reality of speech units in Japanese. In T. Otake & A. Cutler (Eds.), *Phonological structure and language processing: Cross-linguistic studies* (pp. 203-216). Berlin: Mouton de Gruyter.
- Morishima, Y. (2013). Allocation of limited cognitive resources during text comprehension in a second language, *Discourse Processes* 50, 577-597.
- Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, 76, 165– 178.
- Muraishi, S., & Amano, K. (1972). *Yoji no yomi-kaki norryoku [Reading and writing ability of Japanese pre-school children]. (Report No. 45),*

- Tokyo: the National Language Research Institute: Tokyo-shoseki.
- Muraki, M., & Nakaoka, N. (1990). HATSUON to SOKUON: Eigo, chugokugo washa no hatsuon [Syllable final nasal and the first member of geminate consonants: In the case of English native speakers and Chinese native speakers] . In M. Sugito (Ed.), *Koza nihongo to nihongo kyoiku, Dai 3-kan: Nihongo no onsei, on-in (ge) [the Japanese language and its education, Vol. 3, Speech sounds and phonological issues of the Japanese language (No. 2)]*. (pp.139-177). Tokyo: Meiji-shoin.
- Murty, L., Otake, T., & Cutler, A. (2007). Perceptual Tests of Rhythmic Similarity: I. Mora Rhythm. *Language and Speech*, 50, 77-90.
- Muto, T. (1987). Relationship between kana character reading and verbal game playing in Japanese young children. *Paper presented at the 9th Biennial Meeting of the International Society for the Study of Behavioral Development*, Tokyo.
- Nakajo, O. (1989). Nihongo no on-in to akusento [Japanese phonological issues and accents]4. Tokyo: Keiso-shobo.
- Nakamura, M., Kolinsky, R., Spagnoletti, C., & Morais, J. (1998). Phonemic awareness in alphabetically literate Japanese adults: The influence of the first acquired writing system. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, 17, 417-450.
- Nakamura, M. & Kolinsky, R. (2014) Multiple Functional Units in the Preattentive Segmentation of Speech in Japanese: Evidence from word illusions. *Language and Speech*, 57(4), 513-543.
- Nasu, A. (2006). Nihongo onomatopoe no gokeisei to on-in koozoo [Word formation and phonological structure of Japanese mimetics]. *Phonological Studies*, 9, 157-164.
- Navon, D. (1986). On determining what is unconscious and what is perception. *The Behavioral and Brain Sciences*, 9, 44-45.
- Nazzi, T., Bertoncini, J., & Mehler, J. (1998) Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*. 24(3), 756-766
- Nazzi, T., Iakimova, G., Bertoncini, J., & Frédonie, S. (2006). Early

- segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences. *Journal of Memory and Language*, 54, 283-299.
- Neisser, U. (1967). *Cognitive Psychology*, New York: Appleton-Century-Crofts.
- Nesdale, A. R. & Tunmer, W. E. (1984). The development of metalinguistic awareness: A methodological overview, In W. E. Tunmer, C. Pratt, & M. L. Herriman (Eds.), *Metalinguistic Awareness in Children*. (pp.12-35). Berlin: Springer-Verlag.
- NHK(Nippon Hoso Kyokai) (1998). *Nihongo hatsuon akusento jiten [Dictionary of Japanese pronunciation and accent]*. NHK Hoso Bunka Kenkyujo (Ed.), Tokyo: Nihon Hoso Shuppan Kyokai.
- Nishinuma, Y. (1995). Personal computer environment for Japanese prosody processing. Paper presented at TASL 8th meeting, NLRI, Tokyo.
- Nomoto, K. (1990). Onsei to on-in [Speech sounds and phonological issues], In M. Sugito (Ed.), *Koza nihongo to nihongo kyoiku, Dai 3-kan: Nihongo no onsei, on-in (ge) [the Japanese language and its education, Vol.3: Speech sounds and phonological issues of the Japanese language II]*. (pp.178-197). Tokyo: Meiji-shoin.
- Ohara, A. (Ed.) (1994). *Nihon-no warabe-uta [Nursery songs of Japan]*. Kyoto : Yanagihara Shoten.
- Oishi, N. & Saito, S. (1999). Phonological impairment in developmental language disorders: In case of developmental dyslexics, *The Japan Journal of Logopedics and Phoniatrics*, 40, 378-387.
- Olson, D. (1991). Literacy as metalinguistic activity. In D. R. Olson & N. Torrance (Eds.), *Literacy and orality* (pp. 251-270). Cambridge: Cambridge University Press.
- Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language*, 32, 358-378.
- 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* (pp.

- 183-201). Berlin: Mouton de Gruyter.
- Otake, T., & Imai, Y. (2001). Nihongo wo bogo to suru yooji no onsetsu no kaikoozoo no ninshiki to sono fuhensei [Awareness of subsyllabic structure by Japanese preschool children and its universality. *Phonological Studies*, 4, 81-88.
- Otake, T., Yoneyama, K. Cutler, A., & van der Lugt, A. (1996). The representation of Japanese moraic nasals. *Journal of the Acoustical Society of America*, 100, 3831-3842.
- Pattamadilok, C., Kolinsky, R., Luksaneeyanawin, S., & Morais, J. (2008). Orthographic congruency effects in the suprasegmental domain: Evidence from Thai. *Quarterly Journal of Experimental Psychology*, 61, 1515-1537.
- Peereman, R., Dufour, S., & Burt, J. S. (2009). Orthographic influences in spoken word recognition: The consistency effect in semantic and gender categorization tasks. *Psychonomic Bulletin & Review*, 16, 363-368.
- Perfetti, C. A., Beck, I., Bell, L. C., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first-grade children. *Merrill-Palmer Quarterly*, 33, 283-319.
- Perin, D. (1983). Phonemic segmentation and spelling. *British Journal of Psychology*, 74, 129-144.
- Piaget, J. (1974). *La Prise de conscience [Being aware conscious]*, Presses Universitaires de France, Paris
- Port, R.F., Dalby, J., & O'Dell, M. (1987). Evidence for mora timing in Japanese. *Journal of the Acoustical Society of America*, 81, 1574-1585.
- Posner, M. (1964). Information reduction in the analysis of sequential tasks. *Psychological Review*, 71, 491– 503.
- Poser, W.J. (1990). Evidence for foot structure in Japanese, *Language*, 66, 78-105.
- Pratt, C. & Grieve, R.(1984a). The development of metalinguistic awareness: An introduction. In W. E. Tunmer, C. Pratt, & M. L. Herriman (Eds.), *Metalinguistic Awareness in Children* (pp. 2-11). Berlin: Springer-Verlag.
- Pratt & Grieve, (1984b). Metalinguistic awareness and cognitive

- development. In W. E. Tunmer, C. Pratt, & M. L. Herriman (Eds.), *Metalinguistic Awareness in Children* (pp. 128-143). Berlin: Springer-Verlag.
- Pylyshyn, Z. W. (1981). Psychological explanations and knowledge-dependent processes. *Cognition*, 10, 267-274.
- Pylyshyn, Z. W. & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism, *Spatial Vision*, 3(3), 179-197
- Ramus, F., Nespore, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73, 265-292.
- Rastle, K., McCormick, S. F., Bayliss, L., & Davis, C. J. (2011). Orthography influences the perception and production of speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 1588-1594.
- Read, C., Zhang, Y., Nie, H., & Ding, B. (1986). The ability to manipulate speech sounds depends on knowing alphabetic reading. *Cognition*, 24, 31-44.
- Reynolds, B. (1998). Phonemic Awareness: Is It Language Specific? *Literacy Across Cultures*, September 2/2.
<http://www2.aasa.ac.jp/~dcdycus/LAC98/SEP98/reyno998.htm>
- Ricketts, J., Bishop, D. V., & Nation, K. (2009). Orthographic facilitation in oral vocabulary acquisition. *Quarterly Journal of Experimental Psychology*, 62, 1948-1966.
- Roediger, H. L., III, & McDermott, K. B. (1993). Implicit memory in normal human subjects. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology*, Vol. 8 (pp. 63- 131). Amsterdam: Elsevier.
- Rosenthal, J., & Ehri, L. C. (2008). The mnemonic value of orthography for vocabulary learning. *Journal of Educational Psychology*, 100, 175-191.
- Rozin, P. (1976). The evolution of intelligence and access to the cognitive unconscious. In J. M. Sprague & A. N. Epstein (Eds.), *Progress in psychobiology and physiological psychology*. (pp. 245-280). New York: Academic Press.
- Rubin, H., & Turner, A. (1989). Linguistic awareness skills in grade one

- children in a French immersion setting. *Reading and Writing*, 1, 73-86.
- Ryalls, J. (1996). *A basic introduction to speech perception*, San Diego, CA, Delmar, Cengage Learning.
- Saito, Y. (2003). Gendai nihongo no onsei; bunsetsuon to onsei kigou [Speech sounds of modern Japanese; segments and phonetic alphabets] in Z. Uwano (Ed.), *Onsei, on-in [Phonetics, Phonology]* (pp.1-21). Tokyo : Asakura Publishing.
- Sakamoto, Y. (2005). Can Japanese listeners segment spoken words in Japanese with phonemes ? *Phonological Studies*, 8, 89-96.
- Sasanuma, S., Ito, H., Paterson, K., & Ito, T. (1996). Phonological alexia in Japanese: A case study. *Cognitive Neuropsychology*, 13, 823-848.
- Sato, K. (2001). Nihongo onsei no seisei to onkyoo tokuchoo [Generation of Japanese speech sound and the acoustic features]. In Jôo H. (Ed.), *Nihongo Kyouikugaku Siriizu dai 3-kan, Compuuta Onseigaku [Pedagogy of the Japanese Language Series, Vol. 3, Computer Phonetics]* (pp.47-106). Tokyo: Ohfu.
- Scliar-Cabral, L., Morais, J., Nepomuceno, L., & Kolinsky, R. (1997). The awareness of phonemes: So close-so far away. *International Journal of Psycholinguistics*, 13, 211-240.
- Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 546-554.
- Seki, A., Kassai, K., Uchiyama, H. & Koeda, T. (2008). Reading ability and phonological awareness in Japanese children with dyslexia, *Brain & Development* 30, 179-188.
- Serniclaes, W., Ventura, P., Morais, J. & Kolinsky, R. (2005). Categorical perception of speech sounds in illiterate adults. *Cognition*, 98, B35-B44.
- Shibatani, M. (1990). *The languages of Japan*. Cambridge: Cambridge University Press.
- Shibamoto-Smith, J. S. (1996). Japanese writing. In P. T. Daniels and W. Bright (Eds.) *The World's writing systems*, (pp.209-217). New York: Oxford University Press.
- Shinohara, S. (2000). Default accentuation and foot structure in Japanese:

- Evidence from Japanese adaptations of French words. *Journal of East Asian Linguistics*, 9, 55-96.
- Shirota, S. (1993). *Nihongo no oto : Onseigaku to on-in ron [Sound of Japanese : Phonetics and phonology]*. Kasukabe : Hituzi Syobo.
- Sipra, M. A. (2013). Impact of English orthography on L2 acquisition, *English Language Teaching*, 6(3), 116-124.
- Spagnoletti, C., Morais, J., Alegria, J., & Dominicy, M. (1989). Metaphonological abilities of Japanese children. *Reading and Writing*, 2, 221-244.
- Stephens, M. A. & Leane, S. (2013). Phonological Awareness as Revealed by Spelling : Do Japanese EFL learners process written English phonologically? *Journal of Language and Literature, Faculty of Integrated Arts and Sciences, The University of Tokushima*, 21, pp.69-80.
- Strange, W. and Shafer, V. L. (2008). Speech perception in second language learners. In J. G. Hansen Edwards, M. L. Zampini, *Phonology and Second Language Acquisition* (pp.153-191), Amsterdam: John Benjamins.
- Studdert-Kennedy, M. (1981). The emergence of phonetic structure, *Cognition*, 10, 301-306.
- Sugito, M. (1989). Onsetsu ka haku ka [Syllable or mora?]: Choon, hatsuon, sokuon. In M. Sugito (Ed.), *Koza nihongo to nihongo kyoiku, Dai 2-kan: Nihongo no onsei, on-in (jo) [the Japanese language and its education, Vol.2: Speech sounds and phonological issues of the Japanese language]*. (pp.154-177). Tokyo: Meiji-shoin.
- Takayama, T. (2011). SOKUON [Mora obstruent]. In H. Jôo, T. Fukumori, & Y. Saito (Eds.), *Onseigaku kihon jiten [Basic dictionary of phonetics]* (pp.490-492). Tokyo: Bensei JP.
- Tamaoka, K., & Taft, M. (1994). Hakuha on-in shori no saisho tan-i to nariuruka: Giji gairaigo no goi seigo handan karano kousatsu [Is the smallest unit in phonological processing equivalent to the smallest unit in orthographic processing? Lexical judgments of Katakana non-words]. *Japanese Journal of Psychology*, 65, 377-382.

- Tanaka, S. (1998). Futto-nai ni okeru boin no kikoe to fukugoogo akusento [Vowel sonority in foot and compound accentuation], *Journal of the Phonetic Society of Japan*, 2(1), 50-62.
- Tanaka, S. (1999). Nihongo no onsetsu to yonpaku no tenpureeto: Senryu to puro-yakyuu seien ni okeru 'Ji-amari' no bunseki [Syllable and four-mora templates in Japanese: analyses of 'Ji-amari' in Senryu poems and baseball cheering]. In Onsei-Bunpoo Kenkyukai [Spoken Language Working Group] Eds., *Bunpoo to onsei [Speech and grammar] II* (pp. 261-290). Tokyo: Kuroshio.
- Tanaka, S. (2000). The avoidance of rising pitch within a syllable: Evidence for the superheavy syllable and the independence hierarchy in Japanese. In M. Amano et al. (Eds.) *Synchronic and diachronic studies on language : a festschrift for Dr. Hirozo Nakano, (Linguistics and philology, 19)* (pp. 521-531). Nagoya: Nagoya University.
- Tanaka, S. (2008). *Rizumu, akusento no 'yure' to on-in, keitai kouzou ['Fluctuation' of rhythm/accent and phonological/morphological structure]*. Tokyo: Kuroshio.
- Tanaka-Welty, Y. (2005). Language-based leaning and reading disabilities in Japan, *The Japan Journal of Logopedics and Phoniatrics*, 46, 148-154.
- Tanaka-Welty, Y., Hyodo, T., Ohishi, N., Wise, B. & Snyder, L. (2006), Yomikaki no shutoku ya shogai to on-inshori-nouryoku tono kankei ni tsuite no kentou [Reading disabilities in Japanese children: A preliminary study on phonological deficits], *Japanese Journal of Learning Disabilities*, 15(3), 319-329.
- Tateishi, K. (1989). Theoretical implications of the Japanese musicians' language. *The Proceedings of the Fifteenth West Coast Conference on Formal Linguistics (WCCFL)*, 8, 384-398.
- Tenpenny, P. L. (1995). Abstractionist versus episodic theories of repetition priming and word identification. *Psychonomic Bulletin & Review*, 2, 339– 363.
- Terao, Y. (1992). Bun-sanshutsu kenkyu kara mita onsetsu, mora [Syllable and mora investigated by the study of sentence production]. In S. Haraguchi (Ed.), *Nihongo no mora to onsetsu-kozo ni kansuru sogoteki*

- kenkyu [A comprehensive study on the structure of the Japanese mora and syllable]*. (pp. 62-75). Nagoya: Nanzan University.
- The National Language Research Institute (1990). *Nihongo no boin, shiin, onsetsu: Choon-undo no jikken-onseigakuteki kenkyu [Japanese vowels, consonants, syllables: Experimental phonetics research of articulatory movements]*. (Report No. 100). Tokyo: Shuei-shuppan.
- Treiman, R. (1983). The structure of spoken syllables: evidence from novel word games. *Cognition*, 15, 49-74.
- Treiman, R. (1985). Spelling of stop consonants after /s/ by children and adults. *Applied Psycholinguistics*, 6, 261-282.
- Treiman, R., & Cassar, M. (1997). Can children and adults focus on sound as opposed to spelling in a phoneme counting task. *Developmental Psychology*, 33, 771-780.
- Treiman, R., Goswami, U., & Bruck, M. (1990). Not all nonwords are alike: implications for reading development and theory. *Memory and Cognition*, 18, 559-567.
- Treisman, A. (1993). The perception of features and objects, In A. Baddeley & L. Weiskrantz (Eds.), *Attention: selection, awareness, and control: A tribute to Donald Broadbent* (pp.5-35). Oxford: Oxford University Press.
- Treisman, A., & Paterson, R. (1984). Emergent features, attention, and object perception. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 12-32.
- Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, 14, 107-141.
- Tsujimura, N. (2007). *An Introduction to Japanese Linguistics*. Malden, MA : Blackwell.
- Tunmer W. E. & Bowey, J. A. (1984). Metalinguistic awareness and reading acquisition, In W. E. Tunmer, C. Pratt, & M. L. Herriman (Eds.), *Metalinguistic Awareness in Children*.(pp.12-35). Berlin: Springer-Verlag.
- Tunmer, W. E. & Herriman, M. L. (1984). The development of metalinguistic awareness: A conceptual overview. In W. E. Tunmer, C.

- Pratt, & M. L. Herriman (Eds.), *Metalinguistic Awareness in Children*.(pp.12-35). Berlin: Springer-Verlag.
- Tunmer, W. E., Herriman, M. L. & Nesdale, A. (1988). Metalinguistic abilities and beginning reading. *Reading Research Quarterly*, 23, 134-158.
- Tyler, M. D., & Burnham, D. (2006). Orthographic influences on phoneme deletion response times. *The Quarterly Journal of Experimental Psychology*, 59, 2010-2031.
- Ujihira, A. (1996). Kashoo to mora : Kashi to onpu no musubitsuki wo megutte [Songs and special morae : On the relation between lyrics and music notes], *Nihongo, Nihon-bunka Kenkyuu [Study of Japanese and Japanese Culture]*, 4, 14-25.
- Uwano, Z. (2003). Akusento no taikei to shikumi [The system and structure of accent]. In Z. Uwano (Ed.), *Onsei, On-in [Phonetics, Phonology]* (pp. 61-84). Tokyo: Asakura Publishing.
- Valtin, R. (1984). The development of metalinguistic abilities in children learning to read and write. In J. Downing & R. Valtin (Eds.), *Language awareness and leaning to read* (pp. 207-226). New York: Springer-Verlag.
- Vance, T. (1987). *An Introduction to Japanese phonology*, Albany: State University of New York Press.
- Vance, T. (2008). *The Sounds of Japanese*, Cambridge: Cambridge University Press.
- van Kleeck, A. (1987). Foreward. *Topics in Language Disorders*, 7, vi-vii.
- Ventura, P., Kolinsky, R., Brito-Mendes, C., & Morais, J. (2001). Mental representations of the syllable internal structure are influenced by orthography. *Language and Cognitive Processes*, 16, 393-418.
- Ventura, P., Morais, J., Pattamadilok, C., Kolinsky, R. (2004). The locus of the orthographic consistency effect in auditory word recognition. *Language and Cognitive Processes*, 19, 57-95.
- Verdonschot, R. G., Kiyama, S., Tamaoka, K., Kinoshita, S., La Heij, W., & Schiller, N. O. (2011). The functional unit of Japanese word naming: Evidence from masked priming. *Journal of Experimental Psychology*:

- Learning, Memory, and Cognition*, 37, 1458-1473.
- Vihman, M. M., Velleman, S., & McCune, L. (1994) How abstract is child phonology?; Toward an integration of linguistic and psychological approaches, In M. Yavas (Ed.), *First and second language phonology* (pp.9-44), San Diego: Singular Publishing.
- Wang, W. S.-Y. (1973). The Chinese language. *Scientific American*, 228, 50-63.
- Ward, R. (2001). Visual attention moves no faster than the eyes, In K. Shapiro (Ed.), *The limits of attention; Temporal constraints in human information processing*, (pp.199-216), Oxford: Oxford University Press.
- Wild, M. (2009). Using computer-aided instruction to support the systematic practice of phonological skills in beginning readers. *Journal of Research in Reading*, 32(4), 413-432.
- Yamada, J. & Banks, A. (1994). Evidence for and characteristics of dyslexia among Japanese children, *Annals of Dyslexia*, 44, 105-119.
- Yuzawa, Ma., Sekiguchi, M., & Li S. (2007). Phonological awareness of English among young Japanese children: Learning methods of English suitable for young Japanese children. *Hiroshima University*, 56, 153-160
<http://ir.lib.hiroshima-u.ac.jp/00022821>
- Yuzawa, Mi. & Yuzawa, Ma. (2009). A study of a process of second language phonological acquisition in young children of an international kindergarten: Roles of phonological working memory, phonological awareness of English, and vocabulary knowledge of Japanese. *Human Developmental Research*, 23, 189-200.
- Yuzawa, Mi. & Yuzawa, Ma. (2013). *Nihongobogo-yoji niyoru eigono chikaku, hassei to gakushuu*, [Perception, production, and learning of English speech by young Japanese native children]. Tokyo: Kazama-shobo.
- Yopp, H. K. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly*, 23, 159-177.
- Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory recognition. *Psychonomic Bulletin and Review*, 5, 683-689.

List of Publication

- Nakamura, M., Kolinsky, R., Spagnoletti, C., & Morais, J. (1998). Phonemic awareness in alphabetically literate Japanese adults: The influence of the first acquired writing system. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, 17, 417-450.
- Nakamura, M. & Kolinsky, R. (2014). Multiple Functional Units in the Preattentive Segmentation of Speech in Japanese: Evidence from word illusions. *Language and Speech*, 57(4), 513-543.
- Morais, J., Kolinsky, R., & Nakamura, M. (1996). The psychological reality of speech units in Japanese. In T. Otake & A. Cutler (Eds.), *Phonological structure and language processing: Cross-linguistic studies* (pp. 203-216). Berlin: Mouton de Gruyter.

Notes

Chapter 1

1. According to Crystal (1991), ‘phoneme’ is the minimal unit in the sound system of a language according to traditional phonological theories. The original motivation for the concept stemmed from concern to establish patterns of organizations within the indefinitely large range of sound heard in languages. The phonetic specifications of the sound (or phones) heard in speech, it was realized, contain far more detail than is needed to identify the way languages make contrasts in meaning. The notion of the phoneme allowed linguists to group together sets of phonetically similar phones as variants, or ‘members’, of the same underlying unit. The phones were said to be realizations of the phonemes, and the variants were referred to as allophones of the phonemes. It is further complicated in Japanese that ‘phoneme’ is translated sometimes into onso(音素) and sometimes into on-in(音韻). Even though Hattori (1990) admits those two kinds of translations (p.3, 1990), he defines on-in as a general term in phonology including , <phoneme>, <toneme>, accentual pattern, <stress> (p.25, 1990). According to Nomoto (1990), on-in is a result obtained from abstraction of onsei (speech sounds, a concrete fact of the individual) which has various levels and kinds, for example those concerning sound quality, i.e., vowels and consonants, and those sound quantity, i.e. accents. The same differentiation seems to exist as regards to ‘syllable’, one at the level of onsei (speech sound) and the other at that of on-in. Some researchers especially in the field of engineering treat ‘on-in’ and ‘onso’ together as translation of ‘phoneme’ (e.g. Kashino, 1991). However in this research framework, the author distinguishes on-in from onso, the former as a more general term and superordinate category in which the latter is placed as subordinate category in the frame of phonology and also in psycholinguistics. For example, phonological awareness and phonological units should be translated into ‘on-in ishiki’ (音韻意識) and ‘on-in tan-i’(音韻単位),

respectively. The former according to Hatano and Inagaki(1992) and the other Japanese psychologists and the latter according to Kubozono (1992) and Japanese phonologists.

2. An exception is *yoon*, which are digraphs that correspond to one mora. They include palatalized consonants and are represented in hiragana with kana endings in /i/, such as き (/ki/), plus a small version of one of the three kana や, ゆ, or よ (/ja/ /ju/, or /jo/), for example, /kja/ is written as /ki/ + /ja/, namely, きゃ.
3. There are two basic views on lexical representation: *Abstractionist theories* view the lexicon as a set of ideal, modality-free units, and *episodic theories* assume that groups of detailed traces collectively represent individual words (see Roediger & McDermott, 1993; Tenpenny, 1995). The abstractionist view is prominent in current theories; perception is typically assumed to involve information reduction, which is decoding of specific episodes (tokens) into canonical representations (types; Morton, 1969; Posner, 1964) (cf. Goldinger, 1996, p.1166).

Chapter 2

1. There is also a position to count Japanese phonemes as 13 including /c/ which is realized as [ts] (e.g. Kubo, 2003).
2. I will transcribe vowel length with the sign (:) according to the convention of the international Alphabet.
3. Henceforth, hyphens and periods are used to mark moraic and syllabic boundaries, respectively.
4. Note that Japanese allows both (voiceless) stops and fricatives to contrast in geminacy. Yet, the phonetic nature of sokuon varies depending on the involved consonant: it is realized as silence only with stops (Fukumori, 2010; Fujimura, 2007; Saito, 2003; Takayama, 2011). Also, as singleton fricatives are generally longer than singleton stops in Japanese, the singleton/geminate duration ratios are smaller for fricatives than for stops, which may turn the length contrast less

perceptible for fricatives than for stops (see Kawahara, in press, for a comprehensive review).

5. For example, analyzing the lyric-tune relationship in Japanese songs, Kubozono (1999a) showed that the frequency with which the composers attribute either one or two music notes (or beats) to two morae depends on the sonority of the special mora.
6. These two sets of kana were created by Japanese between the 8th and the 12th century A.D. in order to transcribe the Japanese syllable by simplifying. Hiragana, ‘kana without angles’, grew out of an increasingly simplified set of cursively written kanji used as man’yogana ‘Chinese characters used phonetically to write Japanese’. The hiragana syllabary consists of 46 characters, supplemented with a set of diacritics. Today’s hiragana forms were fixed by the (former) Ministry of Education’s 1900 regulations on standard kana signs and usage. Like hiragana, katakana derived from man’yogana. At the beginning of the Heian Period (794-1192), small script that could be written between characters and/or between lines was needed in order to write down reading of, or exegetic commentary on, Buddhist sutras (cf. Ogawa, 1982: p.481). Katakana, ‘simple, incomplete kana’, were created by taking parts of established man’yogana, sometime but not always the same man’yogana that were the source for the counterpart hiragana syllable. The present forms and conventions of use for katakana were fixed in 1900, at the same time as hiragana (cf. Shibamoto-Smith, 1996).
7. Its spelling rules are those of English for the consonants, and of Italian for the vowels.
8. Standard (or Hepburn) system was originally proposed by the association, “Romaji-kai” (organized by Toyama, et al.) in 1885 employing English for Japanese consonants and Italian for Japanese vowel, while Nihon system was originally proposed by “Nihon Romaji-kai” (by Tanakadate, et al.) on the basis of gojuonzu. Because of the request from several international organizations around 1930, Japanese government tried to unify these two Romaji systems and decided to adopt Nihon system in 1937 after the fierce controversy

between the phonologists with the Prague school theory who supported Nihon system and the phoneticians who supported Standard system. (In the international linguistic congress in 1932, Turbetzkoy evaluated highly Nihon system because it was an good example showing that phonology had a great significance in constituting/creating the practical way of spelling/orthography. Kunrei system is a modified version of Nihon system promulgated by Japanese government in 1954, in which some of the transcriptions of the kana of da-column and wa-column are different; {da, di, du, de, do}, {wa, wi, u, we, wo} in Nihon system, while {da, zi, zu, de, do}, {wa, i, u, e, o} in Kunrei system (cf. Koizumi, 1978).

9. The original gojuonzu (literal meaning: 50-sound-table) is highly evaluated even by modern Japanese phonologists (e.g. Tsukishima, 1964), in which fifty basic morae of Japanese at that time were arranged in systematic way from the phonological point of view according to the structure of Japanese basic mora (/V/ or /CV/ syllable): the morae which share the common initial consonant are arranged vertically in the column called 'gyo'; those which share the common final vowel horizontally in the row called 'dan'. The present Kana syllabary of Modern Japanese is still based on the original gojuonzu of Old Japanese although several minor change are included in the order of vowel, in Kana itself and the sound, and the variation of these fifty morae is added to the original covering only chokuon (/V/ and /CV/ morae) which are seion
10. Some researchers (Nesdale & Tunmer, 1984; Tunmer & Bowey, 1984; Tunmer & Herriman, 1984) prefer to use the term, *Metalinguistic Awareness* instead of metalinguistic abilities and/or linguistic awareness. Although metalinguistic awareness is related in meaning to the term '*matalanguage*', it is important to distinguish between the two. While metalanguage refers to language used to describe language, and includes terms like *phoneme*, *word*, *phrase*, etc., metalinguistic awareness refers to awareness of the *instantiations* of these terms, but not to knowledge of the terms themselves. Thus, a metalinguistically aware child may perform well on a task involving the manipulation of phonemes without

knowing what the term *phoneme* means (cf. Tunmer & Herriman, p.12, 1984).

11. Phonological awareness and metaphonological ability are used in the same meaning in this study.
12. According to Carillo (1994) the central role of phonological awareness in learning to read and write in an alphabetic code is becoming widely recognized. Nevertheless, there have been differences in interpreting its role in reading acquisition. Some authors (i.e. Bradley & Bryant, 1985; Lundberg, Olofsson & Wall, 1980) suggest it as a precursor of reading while other (i.e. Morais, Alegria & Content, 1987) claim that, at least some aspect of phonological awareness results from reading acquisition, or is a concomitant of it (cf. Brady & Shankweiler, 1991)
13. The scores reported here are mean percentages of correct responses whereas Spagnoletti et al. (1989) presented the medians.
14. Phonemic awareness covers ability of conscious analysis of speech unit which manifests more conscious representation than just phonemic “sensitivity”. Endo employs a Japanese term of “sensitivity” in order to describe “awareness”, but Morais argues in his literature (1991, p.44), mentioning their findings about a Portuguese illiterate, that “despite his expertise in rhythm and despite his high ability in task requiring sensitivity to subtle relationship of phonological similarity, F.J.C. is unable to isolate intentionally the phonemes of an utterance and to manipulate them”, the sensitivity should also be distinguished from awareness. Amano (1988) also states that ‘the clear consciousness of phones which constitutes a word is called phonemic awareness and which phonemic analysis is often employed instead of it. Speaking more accurately, however, phonemic awareness is a product of phonemic analysis.’

Chapter 3

1. For more details on the Japanese phonology, see for example Shibatani (1990), Tsujimura (2007), and Vance(2008).
2. There were no differences between simple and yoon items on the frequency of mora- or syllable-based patterns of responses. Consequently, both structures were grouped together.
3. This implied that we did not take position as for the linguistically monohonemic (one palatalized consonant) or bi-phonemic (C followed by the glide /j/) status of the sequence Cj. Besides, for this latter alternative, the reversal of e.g. /opja/ was transcribed as a V₂iCV₁ response (namely, /aipo/) because /Vj/ sequences do not exist in Japanese.
4. There is however an alternative approach in which the palatal sound before /i/ and the alveolar affricate sounds before /a/, /u/, and /o/ are considered as corresponding to the same phoneme, noted /c/. Accordingly, while /ci/ is pronounced [tʃi], /ca/, /cu/ and /co/ should be pronounced [tsa], [tsu], and [so], respectively. The phonetic sequences [tʃa], [tʃu], and [tʃo] are then phonemically analyzed as /cja/, /cju/, and /cjo/, respectively (e.g., Hattori, 1979; Nakajo, 1989, The National Language Research Institute, 1990).
5. This is only true for the *standard* (also called *Hepburn*) system. For some kanas there are two alphabetic spellings, the second one being the *kunrei* (or *Nihon*) system, which is based on the phonemic transcription. But even if the kunrei system is taught at school, it is the standard system which is the most often used in everyday life, and which is referred to in the discussion and in the data analysis.

Chapter 4.

1. Since there are large inter-speaker variability in terms of where lexical accents fall, the target word *zange* can also be HLL (NHK, 1998); however, to match it with the other target word of the pair (*shukko*), we used here its LHH version.

2. In fact, for these participants, two targets (i.e., *soegi* and *moeru*) were written with one kanji and two kanas (e.g., 燃える, with /mo/ in kanji + /e/ in hiragana, and /ru/ also in hiragana), because the last two morae of these words correspond to inflections and hence cannot be written in kanji.
3. As suggested by Macmillan and Creelman (2005), to avoid infinite d' values, proportions of 0 and 1 were adjusted to $1 / (2N)$ and $1 - 1 / (2N)$, respectively, where N is the number of trials on which the proportion is based.
4. For all experiments, we checked that similar results were obtained in the analyses using FA rates.
5. Interestingly, CVJCV (e.g., *raisu*) and CVNCV (e.g., *konte*) words differed in their proportions of “two” and “three” responses. While CVJCV words were associated with 62% and 25% of “two” and “three” responses, respectively, CVNCV led to 79% and 8%, respectively (the proportion of “one” responses was 13% in both cases). The comparison between the arcsine transforms of the proportions of “two” responses for these two item types was significant, $t(7) = 2.45$, $p < .05$. Thus, at least with these Japanese items, it seems that syllabic boundaries are less clear with CVJCV than CVNCV words. These results are consistent with the finding of Kubozono (1999a) that there is a hierarchy of independence among the Japanese special morae (see also Tanaka, 2000; Tanaka, 2008). Coherent with these data, a further segmentation test run on 16 CVJCV words and 9 fresh participants led to 41% and 22% of “two” and “three” responses, respectively ($\chi^2(1) = 7.6$, $p < .01$; “one” responses: 37%).
6. This was done in order to ensure that participants had the same level of exposure to or expertise with foreign languages. We additionally checked that similar results were obtained when considering the three groups of participants (i.e., including also the Japanese participants tested in Belgium with the material from Experiment 1). The interaction between constituent, trial type, and group was also significant, $F(6, 243) = 4.87$, $p < .0001$.

Chapter 5

1. However, because of the insufficient number of the participants in each school grade and in order to focus on the target objectives, those participants were not included in the present research analysis.
2. They cited Journal of the Acoustical Society of America, Journal of Phonetics, and Language and Speech (Derwing & Munro, p. 382, 2005).
3. Nishinuma (1995) developed a CAI system to show the simplified visual image of abstract phonological information on the computer screen, but the system aimed to support learning the suprasegmental issue of prosody.
4. Dyslexics are the persons with dyslexia which is a disorder manifested by difficulty in learning to read despite conventional instruction, adequate intelligence, and socio-cultural opportunity (e.g. Ellis, 1993). Dyslexia is a specific learning disability that is neurobiological in origin. It is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities. These difficulties typically result from a deficit in the phonological component of language (Lyon, Shaywitz, S., & Shaywitz, B., 2003, p.2).
5. Some 45 years ago, Makita (1968) claimed that dyslexia was indeed rare in Japan according to his questionnaire study, showing that the incidence of dyslexic children in Japan was less than 1%.
6. Some research asserts that there is no or less effect of CAI for phonemic segmentation training. For example, Kershield, et al. (1994) examined the different types of phonemic segmentation training on phonemic segmentation, reading and spelling. Children with learning disabilities, who were weak in phonemic segmentation, were trained with the computer-assisted training programs. Their findings show that visual support had no beneficial effects. They concluded that with these children, phonemic segmentation training using visual support does not have any advantage over auditory training alone.
7. Their results, however, further showed awareness of rhymes was not facilitated by either high letter-sound knowledge or print feedback. They

concluded that the new definition of literacy should include proto-literacy – knowledge of letter-sound and other print-sound relationships that are learned before becoming literate and that may influence the acquisition of awareness of some sub-syllabic units of speech.

Appendix

List of Appendix

APPENDIX A. Form of the interview questionnaire for the experiments in Chapter 3 & Chapter 4.

APPENDIX B. Instruction for the experiments in Chapter 3.

APPENDIX C. Texts used in the reading test for evaluating the participants' abilities of English and French languages in Chapter 3.

APPENDIX D. Instruction for the experiments in Chapter 4.

APPENDIX E. A page of the answer booklet of the material presented in hiragana used in Experiment 1 in Chapter 4.

APPENDIX F. A page of the answer booklet of the material presented in kanji used in Experiment 3 in Chapter 4.

APPENDIX G. A page of the answer booklet of the material in katakana used in Experiment 3 in Chapter 4.

APPENDIX A. Form of the interview questionnaire for the experiments in Chapter 3 & Chapter 4.

(in Japanese)

(アンケート)

本日は実験にご協力頂きまして誠にありがとうございました。皆様から頂きました貴重なデータをもとに、日本人の言語認識に関する研究を発展させて頂きたいと存じますが、今後の展開の参考までに、次のアンケートにお答え頂きたくよろしくお願い致します。

尚、アンケートの内容は実験のデータ同様、実験、研究の目的以外には使用しないこと、又、お名前は公表致しませんことをここでお約束致します。

ブリュッセル自由大学実験心理学研究所 中村美代子

- 1) 氏名
- 2) 生年月日
- 3) 現在の職業
- 4) 過去に外国語を含めた言語教育、または外国語を使った職業、又は社会活動（研究、ボランティア等）につかれたことのある場合には、次の質問にお答え下さい。
 - ・ 職業、又は活動（だいたいの内容）
 - ・ 期間
- 5) 「かな」を習った時の記憶
 - ・ 小学校入学以前
誰にどのように習ったか、もし覚えていたら書いて下さい。
 - ・ 小学校入学以降
「あいうえお」の表をどのように習ったか、もし覚えていたら書いて下さい。
色を使って習ったり覚えたりした記憶はありますか。
(はい、いいえ)
- 6) ローマ字を習った時の記憶
 - ・ どのように習ったか、もし覚えていたら書いて下さい。

- ・ ローマ字を習った時の印象を（難しかったなど）もし覚えていたら書いて下さい。

7) 外国語習得歴 （習った外国語と期間を書いて下さい。）

学校で

学校以外のところで

（外国の場合は国名も）

小学校時代 語 年 語 年（ ）

中学校時代 語 年 語 年（ ）

高校時代 語 年 語 年（ ）

高校卒業以降 語 年 語 年（ ）

社会人になってからもあれば書いて下さい（海外在住経験も含め）。

現在、習っている場合はそれも書いて下さい。

以上、ご協力ありがとうございました。

English translation by myself

[QUESTIONNAIRE]

Thank you very much for your participation in the experiment. We will explore the investigation of language recognition by Japanese people on the basis of these precious data that you provided us. For the future development of our research, we would like you to fill out the following questionnaire. We promise you that the content of the questionnaire will never be used for the other purpose than our research based on this experiment as well as the data of the experiment and your name will never be exposed in public.

February 27, 1991,

Miyoko Nakamura, The Laboratory of Experimental Psychology,
Free University of Brussels

- 1) Name:
- 2) Date of Birth:
- 3) Present Occupation:
- 4) If you have the experience of being engaged in any job or social activity (research, volunteer work, etc.) concerning the language education (including foreign language education) and/or foreign language itself, please answer the following questions.
 - Occupation/Activity (outline):
 - Period
- 5) Your memory of learning 'kana'
 - Before entering into elementary school:
If you remember who instructed you and/or how, please write it down.
 - After entering into elementary school:
If you remember how 'kana matrix'(Gozuon-zu) was instructed, please write it down.
Do you have a memory of learning or memorizing 'kana' by using colors? (Yes. No.)

6) Your memory of learning 'romaji'

- If you remember how 'romaji' was instructed, please write it down.
- If you remember the impression of learning 'romaji', please write it down. (Ex. Difficult, etc.)

7) Your experience of foreign language: Please write down foreign language(s) and the period.)

(Please also indicate the name of the country where you learn it if it is necessary.)

During the period of _____ at school _____ other
institution (if necessary)

Elementary School:

Junior-high School:

High School:

After graduating high school:

Please explain your experience even after you went out into the real world (society). (Indicate also your experience of living in the foreign country.)

If you learn any foreign language at present, please write it down.

Thank you very much for your cooperation.

APPENDIX B. Instruction for the experiments in Chapter 3.

(Japanese instruction with an English translation)

(In Japanese.)

今日は一緒に言葉遊びをしたいと思います。これはテストや検査では決してありませんので、どうかご自宅にいらっしゃるつもりで、言葉を使って遊んで下さい。

最初に例をいくつか聞いて頂きます。次に、例にならって、1つテープの単語を聞くごとに、ご自分でお考えになったように聞こえてきた音をひっくりかえし、声に出して言って下さい。

聞き取れなかったり忘れたりした場合には、もう一度（何回か聞き取れるまで）聞くことができます。答える方のやり直しもできます。例がよくわからなかったり、やり方が難しい場合にも何度でも例を聞くことができますので、おっしゃって下さい。

尚、この遊びは正答を答えて頂くものではありません。

では、始めますので、よろしくお願いします。

(English translation by myself.)

Today, I would like to play a “word game” with you. This is not a test nor examination, so please make yourself comfortable, and play with the words and manipulate them.

First, please listen to some examples. One example consists of a pair of items, a given item and a response. Then the game item will be presented on the tape recorder. Please listen to it carefully, and say aloud what you think is an answer in the same way as in the examples.

When you can't hear or remember the item, you can listen to it once again or several times until you can recognize it. You can also try the answer once again if you want to. When you don't understand the examples or you have difficulty with the game, you can listen to the examples several times until you can understand them. In such cases, please tell me.

This is just a game and does not require any correct answer.

Thank you very much in advance. Now let's begin.

APPENDIX C. Texts used in the reading test for evaluating the participants' abilities of English and French languages in Chapter 3.

C-1. English text

The transformation certainly showed no sign of slowing yesterday as Forget reached the quarter-finals of the Stuttgart Classis with a 7-6, 7-6 victory over Jakob Hlasek, whose own form has shown a marked upturn in recent months.

The fact that Hlasek is a long-time friend and the Frenchman's doubles and surfing partner made the victory rather more bitter and probably a little less emphatic than it might have been because, in his present mood, Forget has the power of serve and the all-court ability to demolish all but the very best.

(Extracted from THE TIMES, February 22, 1991.)

C-2. French text

Guy Forget continue de suprendre agréablement: le numéro un français n'a certes pas gagné le tournoi de Stuttgart auquel il participait la semaine dernière, mais il a atteint les demi-finales après avoir battu d'excellents joueurs comme Auilera, Hlasek (son partenaire de double) et Ivanisevic, et il ne s'est incliné, samedi, que devant le Suédois Svensson. Et d'extrême justesse, car le sort de la partie n'a basculé que sur deux ou trois coups: Forget a enlevé le premier set assez facilement (6-2), et n'a perdu le deuxième que de très peu au tie-break (8 points à 6) après avoir eu une balle de match!

(Extracted from LE FIGARO, February 25, 1991.)

APPENDIX D. Instruction for the experiments in Chapter 4.

[実験教示]

これからヘッドフォンをつけて頂きます。ヘッドフォンは皆様の頭に合わせて調節できますので、つけてから自由に合わせて下さい。

テープがスタートしますと、両方の耳から同時に違うものが聞こえてきますが、聞こえてくるのは実際にある言葉だったり、実際にはない言葉だったりします。(実際にはない言葉というのは日本語で発音できる文字の組み合わせですが、実際にある日本語ではないということです。) お手元の回答用紙には、実際にある言葉が並べられています。1 ページにつき縦の欄が 2 つあり、それぞれの欄に 30 項目あります。

皆様をお願いすることは、回答用紙に従いながら、書いてある言葉が聞こえたかどうか (どちらの耳からでも結構ですので)、聞こえた場合には+の印を、聞こえなかった場合には-の印を、言葉の横の空欄に書いて頂くというものです (聞こえたような気がする、聞こえなかったような気がするでも同様に結構です)。両方の耳から聞こえてくる最初の 1 組が最初に書いてある言葉かどうか、次の組が次に書いてある言葉かどうかという具合にして進んでいきます。尚、1 組聞いて頂いた後、次の 1 組までの間は 2 秒間です。

特にお願いしたいのは、答えがよくわからなくても、迷っても、多少の誤差は関係ありませんので、とにかくどの項目もぬかさずに(わからない場合でも適当に) 答えて頂きたいということです。そうしませんと、そのあとの順番がすべて狂い、間違ったデータになってしまうからです。途中でどれを聞いているかわからなくなってしまった場合には、慌てずにすぐに手を挙げて下さい。こちらでテープを止め、切れ目のあるところまで巻き戻して再開して頂きます。その間 (また、2 人一緒に聞いて頂く場合、手を挙げなかった方も)、すでに印をつけたところはそのままにしておいて下さい。

1 ページ終わるたびに、ページをめくる為に 12 秒間休みをとってあります。結構長いので 2 ページ一緒にめくらないよう、あせらずゆっくりとめくって下さい。最初のページは練習用です。練習が終わったところで一旦テープを止めおわかりになったか確認します。この時点でわからなところがあれば遠慮なく聞いて下さい。また、聞こえてくる音量が丁度良いかどうかもおっしゃって下さい。

テープは約 35 分間、ページ数にして 11 ページあります。大変長くてお疲れになりますので、その間に皆様の状況によって短い休憩を入れますからヘッドフォンをはずしてお休み下さい。それではよろしく願います。

Please wear the headset on your ears. The headset can be adjusted to your head.

When the audiotape starts playing, you can hear the different sounds, which will be either real words or non-existent pseudowords, in both your ears simultaneously. (“Non-existent pseudowords” means some combinations of letters that can be pronounced as Japanese sounds, but don’t constitute any existent Japanese words.)

In the answer sheets of the booklet on the desk, all the items really are existent Japanese words. On one sheet, there are two columns, each of which includes 30 items. Using the pencil provided, please place a mark in the blank just beside the written item as follows: + when you hear the item in the list, and – when you don’t hear the item. (If you don’t know or don’t feel that you heard it, please skip that item and continue with the next one.) You should judge whether the first pair you hear from your different ears simultaneously is the item in the first cell, and then, judge whether the second pair is that in the second cell, and so on, as you proceed. You have two seconds between one pair and the succeeding pair.

I would like you to do a special favor, that is, even if you don’t have any confidence in your answer, please answer all the items (if you don’t figure out the answer, choose the answer at random) and avoid missing any items. Otherwise, the heard items will not match the written items, which will result in erroneous data. If you get lost and can’t find which item you hear, please raise your hand. I will stop playing and rewind to the appropriate point and you can start again from that point. During that time, leave the mark you have already made as it is. (In the case that two persons hear the tape together, I would like to ask the other person who doesn’t raise their hand to do the same as well.)

In order to turn the page, there is a very short break of 12 seconds. You have enough time, so please don’t rush, and avoid turning two pages together. The first page of the booklet is for practice. After you hear all the practice items, I will confirm whether you understand what to do. If you have any questions, please don’t hesitate to ask me. I also would like you to tell me whether the volume is at the right level for you.

You will hear the audiotape for about 35 minutes and have eleven pages of answer sheet in the booklet. Some of you will find it very long and tiring, so please let me know when you get tired. I will give you a short break and you can take off your headset and just relax. Thank you very much in advance. (My translation)

APPENDIX E. A page of the answer booklet of the material presented in hiragana used in Experiment 1 in Chapter 4.

1	ざんげ	
2	もえる	
3	げんど	
4	ざんげ	
5	まいど	
6	はいる	
7	しゅっせ	
8	まいど	
9	しゅっせ	
10	げんど	
11	しゅっか	
12	はいる	
13	しゅっせ	
14	しゅっこ	
15	もえる	
16	はいる	
17	まいど	
18	しゅっか	
19	そえぎ	
20	はいる	
21	しゅっか	
22	げんど	
23	しゅっこ	
24	だっぴ	
25	まいど	
26	しゅっか	
27	もえる	
28	だっぴ	
29	もえる	
30	しゅっせ	

31	しゅっか	
32	はいる	
33	まいど	
34	しゅっこ	
35	もえる	
36	げんど	
37	だっぴ	
38	まいど	
39	入る	
40	げんど	
41	そえぎ	
42	しゅっせ	
43	まいど	
44	ざんげ	
45	しゅっせ	
46	もえる	
47	しゅっか	
48	そえぎ	
49	もえる	
50	しゅっこ	
51	そえぎ	
52	はいる	
53	もえる	
54	まいど	
55	げんど	
56	もえる	
57	しゅっこ	
58	もえる	
59	しゅっせ	
60	もえる	

**APPENDIX F. A page of the answer booklet of the material
presented in kanji used in Experiment 3 in Chapter 4.**

1	懺悔	
2	燃える	
3	限度	
4	懺悔	
5	毎度	
6	入る	
7	出世	
8	毎度	
9	出世	
10	限度	
11	出火	
12	入る	
13	出世	
14	出庫	
15	燃える	
16	入る	
17	毎度	
18	出火	
19	添え木	
20	入る	
21	出火	
22	限度	
23	出庫	
24	脱皮	
25	毎度	
26	出火	
27	燃える	
28	脱皮	
29	燃える	
30	出世	

31	出火	
32	入る	
33	毎度	
34	出庫	
35	燃える	
36	限度	
37	脱皮	
38	毎度	
39	入る	
40	限度	
41	添え木	
42	出火	
43	毎度	
44	懺悔	
45	出世	
46	燃える	
47	出火	
48	添え木	
49	燃える	
50	出庫	
51	添え木	
52	入る	
53	燃える	
54	毎度	
55	限度	
56	燃える	
57	出庫	
58	燃える	
59	出世	
60	燃える	

**APPENDIX G. A page of the answer booklet of the material in
katakana used in Experiment 3 in Chapter 4.**

1	トイレ	
2	ダンテ	
3	パンツ	
4	ボイル	
5	トイレ	
6	ライス	
7	コンテ	
8	ダンテ	
9	ボイル	
10	トイレ	
11	ライス	
12	パンツ	
13	コンテ	
14	ボイル	
15	ダンテ	
16	トイレ	
17	ライス	
18	パンツ	
19	ボイル	
20	コンテ	
21	ダンテ	
22	パンツ	
23	ボイル	
24	ライス	
25	コンテ	
26	ダンテ	
27	トイレ	
28	ボイル	
29	ライス	
30	パンツ	

31	コンテ	
32	ボイル	
33	パンツ	
34	ダンテ	
35	ライス	
36	パンツ	
37	ボイル	
38	トイレ	
39	ダンテ	
40	コンテ	
41	パンツ	
42	ライス	
43	ボイル	
44	ダンテ	
45	トイレ	
46	コンテ	
47	パンツ	
48	ダンテ	
49	ボイル	
50	トイレ	
51	コンテ	
52	パンツ	
53	コンテ	
54	ダンテ	
55	ライス	
56	ボイル	
57	トイレ	
58	パンツ	
59	ダンテ	
60	トイレ	