VARIATION IN THE USE OF INNOVATIVE KATAKANA IN A JAPANESE CORPUS

by

© Poo Ching FENG

A thesis submitted to the

School of Graduate Studies

In partial fulfillment of the requirements for the degree of

Master of Arts

Department of Linguistics

Memorial University of Newfoundland

July 2018

St. John's, Newfoundland and Labrador

ABSTRACT

This study assesses writer's use of innovative katakana forms in texts in a corpus of written Japanese and examines the effects of linguistic and social factors on the use of innovative katakana. Focusing specifically on innovative katakana that represent sequences including /w/ and /v/ phones, one of the aims of the study is to investigate whether the presence of /w/ in the native phonological system encourages the use of innovative /w/ forms in sequences that do not appear in native lexical items. This is in contrast to forms containing /v/ which lack native counterparts in any context in the Japanese phonological system. Another objective of this paper is to investigate whether the likelihood of using innovative katakana is affected by position within a word. Also, this research applied the framework of variationist sociolinguistics to identify which social factors significantly affect the innovative writing behavior.

To answers these questions, this study uses the data collected from the Chunagon database for descriptive analysis and multivariate analysis. The Chunagon corpus is a written corpus compiled by the National Institution of Japanese Language and Linguistics which contains approximately 100 million words and includes texts published between the 1970's and the 2000's. The results show that the presence of native phones triggers a higher usage of innovative katakana in loanword forms containing /w/ as compared to /v/ forms. The findings also indicate that innovative /w/ and /v/ forms occurred more often in word-initial position than in medial and final positions. Concerning the social factors, the multivariate results show that there is no effect for age and gender for the /w/ variable but for the /v/ variable, there is an effect of gender. Only the /v/ variable showed expected innovative preference in informal (webs and books) genres whereas the /w/ variable showed an unanticipated higher innovative usage in a formal register which comprises governmental and legal genres. In sum, this study has presented some significant findings of showing linguistic reasons (the presence of native phone and the positional effect) that would enhance the use of innovative katakana in written texts.

Keywords: Japanese, katakana, innovative, conservative, corpus, orthography, linguistic factors, variationist sociolinguistics

ACKNOWLEDGEMENTS

Foremost, I would like to express my most profound appreciation to my supervisors, Dr. Paul De Decker and Dr. Sara J Mackenzie whose contribution in stimulating suggestions and encouragement. Both of them patiently guided me throughout this study, especially in writing this thesis. I am grateful to the professors of Departments of Linguistics, who had (and had not) taught me, for providing support and assistance to me. My sincere thanks also go to Daniel Johnson, for suggesting bar graph plotting function in R which helped me a lot in presenting the visualized graphs.

My completion of this thesis could not have been accomplished without the support of my linguistics colleagues, especially sociolinguistics gang members: Mariana Hernandez and Leila Ziamajidi with whom I share my thoughts and feelings.

Last but not the least, I would like to thank my family whose love support are with me in whatever I pursue. Most importantly, my heartfelt thanks to my loving and supportive husband who provides endless inspiration and encouragement while I completed my work.

Table of contents

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1 Background and Motivation of the Study

The Japanese language, unlike other languages, has multiple different writing scripts in its orthography: kanji, hiragana, katakana, and another two possible types of romaji (the Latin alphabet) and the English (Latin) alphabet (Stanlaw, 2002). In general, kanji is used for integrated Chinese loanwords; hiragana is the script for native Japanese words and grammatical components in the language; and katakana is predominantly used to represent foreign words (except loans from Chinese) and loanwords in Japanese (Hogan, 2003; Kay, 1995). As a second language learner to Japanese, I have always been interested in the katakana scripts due to its main function of representing foreignness. Most of the time, katakana characters are used to transliterate non-Chinese donor words so that the loanwords sound like the donor sounds. For example, katakana タクシー/takuʃi:/ for "taxi", katakana プレゼント/purezento/ for "present", katakana ケーキ /ke:ki/ for "cake", katakana ポケット/poketto/ or ポッケ/pokke/ for "pocket" and so on. Sometimes, these loanwords can easily back-transliterate to the original words in the donor language. However, due to Japanese phonemic and phonotactics constraints, some loanwords have to use the traditional katakana forms to replace the donor sounds, resulting in loanword adaptations which are pronounced quite differently from their donor sounds. For instance, katakana パーチー /pa:ʧi:/ is used for "party", katakana ビデオ/bideo/ for "video", ラジオ/razio/ for "radio", etc. These kinds of adaptations made the loanwords sound unlike the donor words because the donor consonants are integrated by using Japanese consonants, such as /ʧ/, /b/, and /z/ in the examples above. This type of inaccurate adaptation has confused me since the time I started learning

Japanese. This leads me to support the view from Coulmas (2003, p. 81) that katakanized words have expanded the difficulty in predicting the equivalent donor inputs.

Fortunately, due to an extraordinary number of English-based borrowings in Japanese (Irwin, 2011), several sets of novel characters have been extended in the katakana writing system since the 1950's (Stanlaw, 2002). As a result, a series of orthographic reforms have occurred to satisfy the need for distinctive characters to accommodate the new foreign sounds that have been adopted into Japanese. These novel katakana characters are derived from the traditional katakana characters as shown in the section 2.4 or as recorded in Stanlaw's (2002, p. 569) innovative katakana chart (Appendix C). Donor words like "party", "video", and "radio" now have more than one katakana form in Japanese, i.e., パーティー/pa:ti:/, ヴィデオ/video/, and ラディオ[1]/radio/. The latter forms, arguably, sound more similar to those donor words than the traditional forms, and I, as a second language learner embrace the innovative forms rather than the conservative ones.

However, the emergence of innovative katakana does not eliminate the use of conservative katakana forms. Instead, both katakana forms coexist in contemporary Japanese. This has attracted attention from many scholars to study both katakana varieties and has motivated me to initiate this study.

Crawford (2007 & 2008) examined the evolution of conservative /ʧi/ to innovative /ti/; Stanlaw (2002) studied on the prevalence of choosing conservative or innovative characters of /tsu/ vs. /tu/, /fui/ vs. /fi/, and /hon/ vs. /fon/ in Japanese dictionaries. These works have inspired me to examine other innovative katakana variants. To be specific, I look at the linguistic factors

[1] Katakana characters are shown in these examples. Elsewhere in the paper, I use mostly IPA symbols to represent those loanwords in a way that is accessible for readers unfamiliar with Japanese scripts. Katakana characters are included when necessary to demonstrate differences between conservative and innovative katakana forms.

that influence the use of innovative and conservative forms in the katakana representing English /v/ and /w/. I detail the linguistic research questions in the following section.

Several previous studies also encourage me to examine social factors that might determine the preference for using innovative or conservative forms in Japanese. For example, many studies (Rebuck, 2002; Dougill, 2008; Barrs, 2011; Inagawa, 2015) show that professional groups, such as the media, utilize more innovative katakana in their productions. A survey by BKK (2012) also indicates that young generations with some English proficiency, use more new katakana words than older age groups. Furthermore, Sakamoto's (2002) study on the frequencies of using traditional /b/ forms and innovative /v/ forms for "violin" and "Beethoven", have further persuaded me to study the same innovative katakana /v/ form.

In addition to the novel labiodental fricative, /v/ forms, I examine the innovative variants of the glide, /w/ in Japanese. The main reason to support the selection of innovative /v/ and /w/ forms for this study is as I noticed from the innovative katakana chart provided by Stanlaw (2002, p. 569), as shown in Appendix C, some innovative /w/ forms were integrated earlier than innovative /v/ forms. So, I conceive a linguistic assumption that might affect the degree of innovative katakana integration in Japanese. I suspect the presence of native sequences in the traditional Japanese system encourages earlier integration of the "missing" sequence(s) as shown in Table 2.5. From here, I further form a hypothesis that the presence of the native sequences will also enhance the frequency of usage of those "missing" (or innovative) sequences in the Japanese language as described in the following chapter.

1.2 Research Questions

In this thesis, I examine the usage of katakana characters by adopting a Labovian variationist framework (Labov, 1972). A writer has two choices for transcribing each English sound: for /w/ and /v/ forms, they can choose either the traditional or innovative character. Is the choice a case of "free variation" or is it sociolinguistically constrained? To answer this, I assess the likelihood of the usage of innovative katakana characters, specifically novel characters for the glide, /w/, and the labiodental fricative, /v/, in written Japanese. I source data from one of the interfaces for NINJAL (National Institution for Japanese Language and Linguistics) corpora, Chunagon corpus.

In the first part, I investigate language-internal factors, namely the role and impact of the presence of native sequences in the Japanese phonological inventory, and the effect of position (where the character is located in the loanword) on the innovative use of /w/ and /v/ segments in loanwords. Accordingly, the linguistic research questions for this study are as follows.

1. Which katakana forms (the innovative or conservative variants) do writers tend to use in their own writing behaviour?
2. Does the presence of similar native sequences affect the appearance of innovative katakana characters in the corpus?
3. In which position relative to word boundaries do the innovative /w/ and /v/ variants occur more frequently in the corpus?

Secondly, I study the social influences (Labov, 2001) of gender and age of the scribes, register types and the diachronic variation through the publication year of the sampling data. I will also examine the effects of katakana recommendations by Japanese authorities (Monbusho, 1991; Bunkacho, 1991; NINJAL, 2003a & 2003b & 2004 & 2006; JTCA, 2015) on the authors' writing behaviour. The followings are the sociolinguistic research questions.

4. Do male writers tend to use more innovative variants than female writers? (regarding the literature from Trudgill (2000), Labov (2001), and Tagliamonte (2011))
5. Do young writers tend to use more innovative variants than older writers? (based on the findings from Rebuck (2002), Dougill (2008), Tagliamonte (2011), Barrs (2011), Irwin (2011), BKK (2012), Inagawa (2015), and Kubozono (2015a & 2015b))
6. Which register contains more innovative variants?
7. Does the use of innovative variants increase over time?
8. What is the extent of the influence of the Japanese bureaucracies' orthographical guidelines?

1.3 Definitions of "variable(s)" and "variant(s)"

As seen in the research questions above, as well as in the following chapters, this study uses the terms of variable(s) and variant(s). So, I provide the definitions of variable(s) and variant(s) here, in order to clarify the differences between these terms.

Referring to Labov (1972), a sociolinguistic variable can be defined as a linguistic variable which has at least two "highly stratified" ways of being expressed (two variants), and the use of either of the variants is influenced by the "age levels or other ordered strata of the society" (p. 8). The variable(s) in this study is considered as a sociolinguistic variable because the corresponding variant preferences of Japanese writers are presumably highly correlated with social background of the speakers and writers as well as with change over time. In this study, the variable is contrasted with its variants, as decribed below:

- "variable(s)": referring to the /w/ and /v/ sequences from the adapted English word
- "variant(s)": referring to the conservative (prestige[2]) katakana form and the innovative (less prestige) katakana form used to render these sequences in Japanese orthography

For example, the word "web" /wɛb/, the variable is /wɛ/ (English) and the variants are the conservative form of ウエ[ue] and the innovative form of ウェ[we] (Japanese).

1.4 The Significance of the Study

As mentioned above and detailed in the following chapter, there are numerous studies of katakana functions and usage, which specifically emphasize the aspects of recognition, comprehension, and frequency of use of distinctive loanwords (Kay, 1995; Rebuck, 2002; Olah, 2007; Daulton, 2004 & 2008; Dougill, 2008; BKK, 2010 & 2012; Inagawa, 2015; Daulton, 2015). Their focal points are rarely on the variation of a particular character or sound unit existing in loanwords as in the studies by Sakamoto (2002), Crawford (2007 & 2008), and Irwin (2011). Therefore, my study is necessary to provide novel data and insights on how likely Japanese writers are to choose innovative katakana, as opposed to conservative ones. Another important aspect of this research is to contribute new perspectives on how linguistic factors, such as the presence or absence of the similar native sequences and the position of sound sequences within words, impact the adaptation and usage of innovative katakana in Japanese.

The findings of this paper are also significant in describing and quantifying variation in use of katakana which occurs in written Japanese. Finally, this study offers some resources for those reviewing the katakana guidelines for non-Chinese loanwords in the future.

[2] The definition of "prestige" is described in the literature review (section 2.5.2).

1.5 Organization of the Study

This paper consists of six chapters including this introduction. The following chapter is the evaluation of past literature on this topic. The third chapter describes the Japanese corpus used in this research and outlines the methods of collecting and analyzing data. Then, I present the descriptive results and multivariate statistics in the Results and Analysis chapter. The fifth chapter is the interpretation and discussion of the findings. I also compare my findings to other relevant previous studies and evaluate the importance of the findings of this study. In the final chapter, I summarize the interpretations as well as point out the limitations of this study. Lastly, I give some suggestions for future directions in studying the reform of innovative katakana in Japanese.

CHAPTER 2: LITERATURE REVIEW

The main purpose of this chapter is to summarize and evaluate previous literature related to variation, orthography, and loanwords in Japanese. The first section introduces the complex orthographic system of Japanese which includes at least three different scripts: two types of kana characters, namely, cursive hiragana (ひらがな) and angular katakana (カタカナ), and another type of kanji which is also referred to as Chinese characters. Section 2.2 provides background on the Japanese phonological inventory as well as the phonological restrictions of the language. Section 2.3 provides a brief review of loanword adaptations in Japanese. In section 2.4, I define innovative katakana and provide a brief summary and evaluation of previous studies which have examined the use of innovative katakana. In the context of previous studies, I also explain the motivation for choosing specific katakana characters as the variables to investigate in this study. In section 2.5, the framework of variationist sociolinguistics as well as the significance of using corpus data in sociolinguistics research is reviewed.

2.1 Japanese Orthography

Japanese orthography is complex as it consists of at least three types of writing scripts in contemporary Japanese. In general, kanji, hiragana, and katakana are the core characters in the Japanese writing system. Kanji are Chinese characters which represent morphemes. Each kanji represents not only a native Japanese morpheme but also a loaned Chinese morpheme. Accordingly, kanji usually have at least two different readings in Japanese; one is a native reading, and the other is a Chinese-derived pronunciation. Hiragana and katakana[3] are categorized as moraic characters

[3]The full version of hiragana and katakana characters are in Appendices A and B.

in the Japanese phonological system.

Table 2.1 provides the category of orthographies used in an example of Japanese writing which is Romanised and glossed in English. As noted in (2.1), Japanese is written in a mixture of kanji, hiragana, and katakana. Romaji (the Latin alphabet) is also commonly used for the transliteration of kanji, hiragana, and katakana.

(2.1) 私はマレーシア人です。 (I am Malaysian.)

Table 2.1: Triplets writing styles in modern Japanese

Japanese	私	は	マレーシア	人	です	。
Romaji	*watashi*	*wa*	*Mare:shia*	*jin*	*desu*	。
English	I	(as for)	Malaysia	person	(to be)	.
Orthography	kanji	hiragana	katakana	kanji	hiragana	(period)

Most literature confirm that Japanese orthography has three types of writing in modern Japanese. However, Stanlaw (2002) claims that the language has an additional two types of writing, i.e., romanization (or romaji in Japanese) and characters of the English (Latin) alphabet. Stanlaw's (2002) claim that romaji and the English (Latin) alphabet are two different scripts is confusing but conceivable even though romaji also uses the same characters as the Latin alphabet. His argument can be interpreted using the example of the English word "hotel". If "hotel" is written in Roman script (without any adaptation process), then it is considered as another type of script, namely the English (Latin) alphabet, in Japanese. But if it is Japanized as /hoteru/ without using katakana characters, then it is considered as romaji (Romanised word): *hoteru*. The use of Roman script in Japanese is also evaluated by Irwin (2011) with more detailed examples showing that Roman script in contemporary Japanese is broadly used "to write acronyms, to write abbreviations, and to write full words" (p. 186).

Stanlaw's (2002) claim that these two types of writings are distinct in Japanese is

unpersuasive due to the role of external reasons for extending these alphabetical scripts into Japanese writings, such as attracting the readers' attention, creating an atmosphere of globalization, using some words in English which are just simpler than existing Japanese words (Ishiwata, 2001) and as decorative English characters on media signs even with spelling errors or ungrammatical structures (Shaad, 1987; Dougill, 2008).

2.1.1 Usage of Different Types of Writing in Japanese

Each of the scripts described above has specific functions in Japanese texts. Kanji characters are particularly used for Sino-Japanese (loans from Chinese characters integrated concretely in Japanese), while hiragana characters usually represent the native grammatical portions in a phrase. For example, the hiragana character of は/wa/ in (2.1) is one of the Japanese particles which represents a topic marker for “I” in this case. In the same example, the hiragana characters of です/desu/, which are used to predicate “Malaysian” in (2.1), indicate one of the Japanese copulas. Katakana script is representative of non-nativeness and is used for loanwords from other languages, whereas, Roman letters are predominantly utilized by non-Japanese to read kanji and the other two moraic scripts (hiragana and katakana).

Compared to other scripts, the katakana script has broader and more innovative functions, as stated by Igarashi (2008). Igarashi (2008) further contends that the role of katakana characters is changing in the Japanese writing system. The katakana script is widely used in gairaigo (loanwords), Sino-Japanese words, mixed-words or in forms which include a mixture of katakana and other script types (e.g., 窓ガラス /mado garasu/ which means “window pane”), onomatopoeia, proper nouns (e.g., Canada /kanada/), and native Japanese words. Additionally, there are other motivations for using katakana loanwords in mass media, such as to truly reflect the events being

reported[4]; to mark emphasis or points of interest; to mark slang and newly created words; to demonstrate trendiness; and to alleviate seriousness or harshness in communication (Wong, 1992).

Katakana characters are noticeably correlated with foreign borrowings in Japanese as a tool in nativizing foreign loanwords. This is because Japanese speakers and writers always Japanize non-native words[5] using katakana characters. This does not mean that Japanese speakers never engage in code-switching. Rather code-switching only occurs under specific circumstances, particularly in English language classes (Olah, 2007; Daulton, 2008 & 2015; Hobbs et al.,2010; Barrs, 2011; Champ, 2014; Inagawa, 2014), amongst bilingual children in families with one non-native Japanese parent (Fotos, 1990; Nakamura, 2005; Mori et al., 2015), and to gain attention in the mass media (Irwin, 2011; Goldstein, 2011) irrespective of the correctness of the English words (Shaad, 1987; Dougill, 2008).

2.2 The Japanese Phonological System

Japanese orthography is closely connected to its phonological system. Although there is some variation in the phonetic realization of certain sounds in borrowings, the basic vowel and consonant inventories in Tables 2.2 and 2.3 are broadly representative of the contemporary Japanese phonological system.

Japanese has five vowel phonemes in its inventory as illustrated below.

Table 2.2: Japanese vowels

	Front	Back
Close	/i/	/u/
Mid	/e/	/o/
Open	/a/	

[4] Japanese (mass media) are in favour of using katakana characters for events' names or titles, and Sakamoto (2002) indicates that innovative katakana characters are often seen in titles of events, books, etc.

[5] The non-native and foreign terms refer to non-Chinese loanword because all loans from Chinese are retained as Chinese characters (or kanji) in the language.

Meanwhile, there are 15 primary consonants in the Japanese consonant inventory which are shown as in Table 2.3 (adapted from Pinter, 2015). The consonants in parentheses are considered as surface allophones.

Table 2.3: Japanese consonants inventory (Pinter, 2015)

	Bilabial	Labiodental	Alveolar	Alveo-Palatal	Palatal	Velar	Uvular	Glottal
Plosive	p b		t d			k g		
Fricative	(ɸ)	(v)	s z	(ʃ) (ʒ)	(ç)			h
Affricate			(ts) (dz)	(ʧ) (ʤ)				
Approximant			r		j	w		
Nasal	m		n	(ɲ)			ɴ	

2.2.1 Mora and Syllable

Japanese consonants and vowels are organized into prosodic constituents: syllable and mora. A mora is a unit used to measure syllable weight. Cross-linguistically, syllables with two moras are considered as heavy syllables, and syllables with one mora are categorized as light syllables.

In Japanese, the mora is vital in the phonological system. Most moras consist of a vowel or a consonant-vowel sequence. These moras can also function as a syllable on their own. Other moras include vowel length, consonant gemination, and a coda nasal. These moras need to incorporate with another mora to constitute a syllable. The following example illustrates the syllable and mora using Japanese automobile companies' names (Kubozono, 2015a, p. 12). Note that the symbol "." is used to mark syllable boundaries whereas "-" is used to indicate mora boundaries in the examples of (2.2), (2.3), and (2.4). The Japanese characters shown below the Roman scripts of 'toyota', 'nissan', and 'honda' are katakana characters. In the first syllable of "Nissan", the CV mora /ni/ combines with a mora that geminates the /s/ of the following syllable. This gives the syllable /nis/ which contains two moras. The second syllable of this word also consists of two moras, the CV mora /sa/ and the coda nasal. For "Honda", the nasal mora of the

first syllable also needs to attach to the preceding /ho/ mora to constitute a syllable, i.e., /hon/. Thus, coda nasal, vowel length and consonant gemination are categorized as “non-syllabic mora” while other moras consist of (C)V sequences and are considered as “syllabic mora”.

(2.2) Toyota	to.yo.ta (3 syllables)	to-yo-ta (3 moras)
	ト.ヨ.タ	ト-ヨ-タ
(2.3) Nissan	nis.san (2 syllables)	ni-s-sa-n (4 moras)
	ニッ.サン	ニ-ッ-サ-ン
(2.4) Honda	hon.da (2 syllables)	ho-n-da (3 moras)
	ホン.ダ	ホ-ン-ダ

It has been argued that, in Japanese, the mora is more significant than the syllable (Kubozono, 2015a). Mora counting is important in traditional poems and speech rhythms[6], as well as in phonological processes and phonotactic constraints. Also, the mora is extensively employed in the analyses of accent, stress, and rendaku (sequential voicing).

This study will also use the mora to examine innovative katakana. This is essential because, in orthographic forms, each character represents a mora. For instance, the syllables of /nis/, /san/ and /hon/ consist of two katakana characters in a single syllable. When counting moras, the number of moras in “Nissan” is the same as the number of katakana characters (four moras and four katakana characters). This also occurs in “Honda” which has three katakana characters and its number of moras is also three.

2.2.2 Japanese Phonotactics

Coda consonants and onset consonant clusters are prohibited in the native Japanese system (Pinter, 2015; Kubozono, 2015b). Some exceptions occur in the coda which allows nasals and voiceless

[6] Mora is used in the analysis of speech rate in Japanese because some scholars consider each mora as timing unit (NHK, 2015) even though, in fact, not all moras have the same phonetic duration in speech (Warner & Arai, 2001; Kubozono, 2015a).

geminate obstruents word-internally and permits only nasals in word-final positions. Other than these exceptions, native Japanese words are restricted to consonant-vowel (CV) sequences, and not all consonants can pair freely with all possible vowels (Pinter, 2015). Kubozono (2015b) lists the CV sequences which are prohibited in the native phonology. Excluded sequences include coronals followed by high vowels ([si], [ði], [ti], [tu], [di], [du]) and fricatives, affricates, and approximants preceding front or central vowels ([fa], [fi], [tsa], [wi]). Repair of these illegal CV sequences results in allophonic variation in Japanese and adaptation in loanwords from other languages, as discussed below.

2.3 Loanword Adaptation in Japanese

In loanword phonology, most foreign inputs undergo some adaptation in order to obey as many native phonological restrictions as possible before being integrated as loanwords. In general, loanword adaptation in Japanese can be divided into two types of adaptation (Paradis & LaCharité, 2011). The first and the most common process is phonological adaptation which is subject to Japanese phonological constraints. Second is non-phonological adaptations which are related to the effect of the orthography of the borrowing language and the perceptual salience of non-native phones and phone sequences. Both types of adaptations are discussed below.

2.3.1 Phonological Adaptation

When borrowings occur in Japanese, foreign consonants are adjusted to the closest "Japanese consonants with the same (or similar) place, voicing, and manner" (Tsuchida, 1995, p. 147). Of course, vowel adaptation follows the same adjustments as foreign consonants do, involving the alteration of vowels in the donor language to Japanese vowels with similar height and backness (e.g. Irwin, 2011, p. 96).

As mentioned in section 2.2.2, consonant clusters are prohibited in Japanese and vowel epenthesis is the typical solution for altering consonant clusters into legal consonant-vowel sequences in order to conform to this phonotactic restriction. The vowel /u/ is the most common epenthetic vowel used to repair the illegal consonant cluster sequences and closed syllables; followed by vowels /o/ and /i/. Choice of epenthetic vowel is affected by adjacent consonants (Irwin, 2011; Shoji & Shoji, 2014; Kubozono, 2015b; Bălan, 2015). For example, /o/ is typically inserted after the dental stops [t] and [d] whereas /i/ is inserted after the palatal-affricates of [ʧ] and [ʤ] (Shoji & Shoji, 2014; Kubozono, 2015b) with epenthetic /u/ inserted elsewhere[7].

According to Kubozono (2015b), phonological adaptations can also be categorized according to two strategies. The first is preservation of most foreign vowels and alteration of some consonants in order to obey Japanese phonotactics. This pattern is consistent with observations by Tsuchida (1995) and Irwin (2011). For example, foreign [si] is modified as [ʃi]; [ti] and [di] are modified as [ʧi] and [ʤi] respectively, maintaining vowel quality and showing effects of consonant changes that are consistent with phonotactic restrictions of Japanese. In this case, adaptation reflects the fact that [ʧi] and [ʤi] occur as allophones of /t/ and /d/ in the environment of a following high, front vowel, /i/.

The second strategy is vowel epenthesis which is a common process in Japanese loanword adaptation for illegal CV sequences from foreign inputs. This strategy is compatible with the explications by Irwin (2011), Shoji and Shoji (2014), Kubozono (2015b), and Bălan (2015). For instance, /fi/ from “film” is realized as /fu.i/; and /wi/ from “whiskey” is separated into /u/ and /i/[8].

[7] Vowel epenthesis in Japanese loanword adaptation has several exceptions when choosing epenthetic vowels, see examples in Irwin (2011), Shoji and Shoji (2014), Kubozono (2015b), and Bălan (2015).
[8] The example of 'whiskey' from Kubozono (2015b) is not compatible with other previous literature (Irwin, 2011) because 'whiskey' is adapted in Japanese with more than two variants as outputs. This example will be discussed further in section 5.1.1.

The first type of adaptation shows explicitly that the Japanese phonological system lacks CV sequences that are identical to the foreign consonants sequences. Many foreign consonants need to be replaced with the closest Japanese consonant so that the outputs obey Japanese phonotactics. Some distinct English consonants share the same Japanese consonant in the outputs. For example, English [b] and [v] are both realized as Japanese [b]; [θ] and [s] are both rendered as Japanese [ʃ]; [ð] and [z] change to [z] in Japanese. In contrast, English [h] is adapted in Japanese as one of three possible consonants: [h], [ç] and [ɸ] depending on the phonetic context. When the following vowel is /i/, [h] is usually adapted as [ç] whereas when [h] is followed by /u/, it is always integrated as [ɸ] in Japanese (Okada, 1999).

Due to the native phonological constraints which prevent consonant clusters and closed syllables in Japanese, the strategy of vowel insertion is considered as the default adaptation strategy in loanword phonology (Irwin, 2011; Shoji & Shoji, 2014; Kubozono, 2015b) as well as in loanword orthography. For examples, coda [v] in English "love" or [lʌv] is epenthesized with /u/ to achieve an output of /rabu/. Similarly, English "street" undergoes epenthesis of /u/ and /o/ to produce Japanese /sutori:to/.

From the point of view of orthography, the Japanese orthographic system has no equivalent characters to those which represent English consonants such as [v] or [s] or [t]. All characters in Japanese represent a mora, typically a consonant-vowel sequence. As a result, vowel epenthesis is obligatory for nearly all foreign consonant clusters so that the outputs can be presented in appropriate katakana forms.

2.3.2 Non-phonological Adaptations

In light of orthographical adaptation in loanwords, previous literature has shown that orthography

of the donor language can influence loanword adaptation such as English borrowings in French (Vendelin & Peperkamp, 2006). However, Paradis and LaCharité's (2011) study shows that English orthographic influence in French and even in Japanese are not significant. For instance, Paradis and LaCharité (2011, p. 765) found that only approximately 1% of loanwords in Japanese and French (Quebec City French) are influenced orthographically, mainly for English vowel adaptation. Although their results of low probabilities of English orthographic effect indicates that orthographic influence is not as strong as previously argued by Vendelin and Peperkamp (2006), the role of orthographic adaptation is still notable in Japanese if the adaptation is based on pronunciations in dictionaries (Irwin, 2011). According to Irwin (2011), pronunciations in foreign dictionaries are the standard reference for Japanese scholars when importing donor words because the pronunciations that appear in those dictionaries are considered as the most established and accurate sources. Therefore, from the literature by Vendelin and Peperkamp (2006), Paradis and LaCharité (2011), and Irwin (2011), it is undeniable that donor orthography (either spelling of donor words or orthographic pronunciation guides in dictionaries) has some influence in loanword adaptation in Japanese, regardless of the extent of orthographic effect.

Apart from orthographical adaptation, perceptual assimilation is another type of adaptation which is widely attested in loanword studies. Usually, perceptual factors in loanwords adaptation are highly correlated with the English efficiency of Japanese authors, publishers, and bureaucrats (Irwin, 2011; Daulton, 2008). Tamaoka and Miyaoka (2003) further explain that Japanese perception is actually also determined by "their daily experience of exposure to katakana words in print" (p. 69). An exceptional study by Peperkamp et al. (2008) contends that most loanword adaptation in Japanese is initiated by perceptual mapping of the foreign sequences to the closest native sounds. They claim that the perceptual assimilation which occurs during the perception

process has influenced how the Japanese adapt English word-final [n] and French word-final [n] differently: English /n/ is rendered as Japanese nasal /n/ whereas French /n/ is epenthesized with vowel /u/. Of course, their result is refuted by other studies such as Smith (2006 & 2009). Smith (2009) disagrees with Peperkamp et al. (2008), arguing that both perception and orthography can affect the borrowing forms in loanword adaptation. Furthermore, perception is only appealed to as an explanatory factor in loanword adaptation in those cases where epenthesis does not occur (Smith, 2006 & 2009). For instance, the word "pocket" has two katakana variants in Japanese: one undergoes adaptation with epenthetic /o/ after [t], /poketto/; the other variant conforms to Japanese phonotactic via deletion, /pokke/. Smith (2006 & 2009) explains that Japanese perceive word-final /t/ due to the influence of donor orthography and then realize that /t/ requires epenthetic /o/ to conform to Japanese phonotactics. So, the epenthesized /poketto/ is considered as a phonological adaptation which corresponds to English orthography. The latter variant, /pokke/, is claimed to be due to the influence of auditory adaptation which is related to perceptual mapping of the closest sound to the origin sound (Smith, 2006 & 2009).

Following this literature, it is understood that perceptual adaptation would produce loanwords that more precisely resemble their pronunciation in the donor language, such as the deleted variant of /pokke/, relative to loanwords formed via phonological adaptation, the epenthesized /poketto/. Therefore, it could be assumed that perceptual adaptation of loanwords in Japanese could lead to the creation of additional characters in order to assimilate foreign sounds as accurately as possible. Such additional characters have been used in orthographic representations of loanwords and are referred to as innovative katakana.

2.4 Innovative Katakana

Recent loanwords in Japanese use not only traditional katakana characters but also novel characters innovated to represent foreign sound sequences. Because some donor words are not adapted to conform to Japanese phonotactics, but rather include sounds and sound sequences which are absent from native vocabulary, innovative katakana characters have been created in order to represent these novel pronunciations. The need for more katakana variants is arguably due to the perception of some Japanese speakers who like to precisely adapt the donor words (Irwin, 2011).

According to my knowledge, previous literature has not attempted to define innovative katakana. Therefore, the definition of innovative katakana provided in section 2.4.1 is based on interpretation and observations from previous studies of katakana scripts in loanwords.

2.4.1 Definition of Innovative Katakana

The creation of novel katakana characters usually involves the amalgamation of a normal size katakana character and another half-sized katakana character representing vowels (/a, i, u, e, o/) or glides (/ja/, /ju/, /jo/). For instance, in example (2.5), the character テ /te/ and the reduced vowel ィ (or half-sized of イ /i/) are combined to create a novel character for the foreign sound sequence ティ/ti/. /ti/ is not found in native Japanese forms where /t/ is realised as the allophone [ʧ] when the vowel /i/ follows. In the novel katakana, the reduced katakana characters replace the vowels in the combined katakana characters, such that the reduced /i/ takes the place of /e/ in /te/ to render novel /ti/ in (2.5).

(2.5)

テ /te/	+	ィ	=	ティ/ti/
(regular size /te/)		(half-sized of イ /i/)		(novel character /ti/)

These combinations of regular katakana with the half-sized vowels or glides certainly

enhance accuracy in borrowing foreign words into the Japanese language, although the creation might disobey restrictions in the native system (Coulmas, 2003). The differences between innovative katakana and conservative katakana are reviewed in the section below.

2.4.2 Conservative vs. Innovative Katakana

Generally, the adaptation of foreign sounds in Japanese can be explained in terms of bimoraic (conservative katakana) adaptation and monomoraic (innovative katakana) adaptation (Irwin, 2011). The bimoraic katakana forms usually consist of two native sequences written with two katakana characters. The bimoraic adaptations use traditional katakana and adhere to Japanese phonotactics. Therefore, bimoraic forms are referred to here as conservative katakana. However, the configuration of monomoraic forms is different, because the existing vowel in the native mora is substituted by half-sized vowels or glides to produce a novel sequence as described above.

For instance, /f/-vowel sequences borrowed into Japanese have two variants. The bimoraic forms use conservative katakana and conform to the phonotactic restriction of native Japanese forms. The monomoraic forms use innovative katakana and represent sequences not found in traditional Japanese vocabulary (Table 2.4). /f/ is not a Japanese phoneme but /f/ occurs allophonically before /u/ and the sequence of /fu/ shares the same kana character with /hu/. In other words, /h/ and /f/ are in free variation when the following vowel is /u/ and the orthographic form フ/hu/ can be realized as [fu] and [hu]. Other sequences of /f/ followed by a vowel in loanwords cannot be represented using traditional katakana. Rather, these sequences can be rendered as bimoraic sequences consisting of the character for /fu/ followed by an additional character representing the vowel found in the donor language, or these sequences can more accurately follow the donor language pronunciation by using innovative katakana. For other /h/-

vowel sequences, the use of traditional katakana forms of ハ /ha/, ヒ /hi/, へ /he/, and ホ /ho/ is retained.

Table 2.4: Conservative (bimoraic) and innovative (monomoraic) katakana characters of /f/ sequences

/f/	Bimoraic (/fu/+vowel)		Monomoraic (/fu/+ reduced vowel)	
Vowel		**katakana**		**katakana**
ア/a/	fua	フア	fa	ファ
イ/i/	fui	フイ	fi	フィ
ウ/u/	fuu	フー	**fu**	**フ**
エ/e/	fue	フエ	fe	フェ
オ/o/	fuo	フオ	fo	フォ

The representation of monomoraic /f/ with adjacent half-sized vowels (with the exclusion of /u/) is among the initial adaptation of innovative characters during the 1950's (Stanlaw, 2002). The bimoraic /f/ katakana characters are considered as conservative forms because all such characters are rendered by combining the native /fu/ with other native vowels. The /f/ bimoraic forms are realized as two significant sounds of [fu] and the corresponding vowel, such as /fui/ in (2.6).

(2.6) Bimoraic (conservative) adaptation:
/fu/ + /i/ = /fui/ フイ

(2.7) Monomoraic (innovative) adaptation:
/fu/ + half-sized /i/ = /fi/ フィ

The column of monomoraic /f/s are the innovative variants (with the exception of /fu/ which occurs natively) which are used to adapt the foreign /fa/, /fi/, /fe/, and /fo/ precisely. Example (2.7) shows the monomoraic adaptation of /fi/ which is innovated by combining the native /fu/ with the reduced vowel /i/.

Because conservative and innovative characters coexist in the writing system, some loanwords have more than one variant in written Japanese. For example, the English word "feminine" is realized with two possible variants in Japanese loanword orthography: **フエ**ミニン

/**fue**minin/ as the conservative sequence, and フェミニン/**fe**minin/ as the novel sequence.

2.4.3 Previous Studies of Innovative Katakana

The adaptation of foreign consonants into native sequences has gradually become an option in written Japanese as innovative katakana that are able represent foreign sounds more precisely have become available. In recent decades, linguistic studies of Japanese have been extended to the domain of the development and diffusion of innovative katakana forms. In phonological studies, Pinter (2015) thoroughly examined the development of three classes of Japanese consonants in loanwords: sibilant fricatives, non-sibilant fricatives (particularly /h/), and coronal stops. In native Japanese forms, the realization of these consonants is dependent on their phonological context with /s/ being realized as [ʃ] before /i/, and /h/ being realized as [ç] before /i/ and as [ɸ] before /u/, before high vowels, /t/ is relized as [ʦu] or [ʧi]. When conforming to native phonotactics, the foreign consonants of /s/, /h/, and /t/, are always realized as [ʃi], [çi] or [ɸu], and [tsu] or[tʃi], respectively, depending on the vowel context. However, Pinter (2015) shows that innovative katakana allows orthographic representations of /si/, /hi/, /hu/, /ti/, and /tu/ and that the use of these forms is increasing over time, particularly in recent katakana words. Pinter (2015) argues that the evolution of these novel consonant sequences is influenced by perception with articulatory motivations.

Other studies such as Crawford (2008) and Stanlaw (2002) also provide strong evidence for the wider acceptance of innovative katakana among Japanese writers and readers today. Crawford (2008) investigates the establishment of the novel character /ti/ versus the conservative counterpart /tʃi/ in modern Japanese, specifically in loanword phonology. His findings show that recent loanwords tend to use novel /ti/ in loanword adaptations as opposed to older loanwords

which are more likely to use the conservative katakana. His study also indicates that the palatazation of /ti/ in Japanese evolved diachronically from before the 1890's where the foreign /ti/ was only katakanized as /ʧi/; then between 1870 to 1930 where /ti/ was gradually shifted to /ʧi/ or /ti/; finally, after 1930, most foreign borrowings containing /ti/ are rendered as /ti/ in Japanese, especially newer loanwords.

Stanlaw (2002) examines the frequency in opting between conservative characters for /tsu/, /fui/, and /hon/ and innovative characters /tu/, /fi/, and /fon/. His study shows that the innovative /fi/ is most commonly used in orthographic rendering of the loanword /firumu/ "film" as opposed to conservative /fuirumu/. In other words, the frequency of /fui/ in "film" is lower than (<<) /fi/ in his findings, as in (2.8).

	Donor word	**Conservative**		**Innovative**
(2.8)	film	/fuirumu/	<<	**/firumu/**
(2.9a)	telephone	/**terehon**/	>>	/terefon/
(2.9b)	earphone	/**iyahon**/	>>	/iyafon/
(2.9c)	headphone	/**heddohon**/	>>	/heddofon/
(2.10a)	telephone quiz	/terehonkuizu/	<<	**/terefon**kuizu/
(2.10b)	telephone news	/terehonnjuuzu/	<<	**/terefon**njuuzu/
(2.10c)	telephone personal advice	/terehonjinseisoodan/	<<	**/terefon**jinseisoodan/

Also, for loanwords which include the English morpheme "phone", the examples from (2.9a) to (2.9c) refer to the physical "phone", such as "telephone" /terehon/, "earphone" /iyahon/, and "head phone" /heddohon/, are older loanwords. Therefore, the frequencies for conservative forms for (2.9a-c) are higher than (>>) innovative forms of /terefon/, /iyafon/, and /heddofon/.

When the "phone"-words refer to telephone-based tasks, such as the examples in (2.10a-c): "telephone quiz" /terefonkuizu/, "telephone news" /terefonnjuuzu/, "telephone personal advice" /terefonjinseisoodan/, the monomoraic innovative form /fo/[9] is the favoured variant. These

[9] /ho/ is the conservative counterpart for the /fo/ sequence in the loanword system. This does not reflect that /ha/, /hi/, /he/ are also the conservative counterparts for foreign /fa/, /fi/, /fe/ respectively. /fu/ is the exception for this

loanwords are relatively newer than the physical "phone"-based loanwords.

Stanlaw (2002) also analyses other innovative katakana forms diachronically, with data collected from dictionaries from the 1950's to the 2000's as presented in Chart C in Appendix C. Stanlaw's (2002) innovative sequences chart has led me to believe that innovative katakana sequences that represent consonant vowel sequences that are ill-formed in Japanese are more likely to be used in the written representation of loanwords when the same consonant does not occur with a different vowel following. For example, /ʤe/, /ʧe/, /ti/, /di/, /fa/, /fi/, /fo/, /wi/, /we/, and /ʃe/ (Table 2.5) are all sequences which are not found in native Japanese and must be represented with innovative katakana if representations which more precisely mirror the donor language pronunciation are desired. However, the sequences recorded in the "Native sequence(s)" columns in Table 2.5 do occur as native sequences. This differs from the case of forms such as /va/, /vi/, /vu/, /ve/, and /vo/, where /v/ does not occur with any following vowel in Japanese, not even allophonically.

Pinter's (2015) findings on innovative consonants /h/ and /t/ and Crawford's (2008) findings regarding /ti/ usage, have led me to expect that the presence of "Native sequence(s)" in Table 2.5 that includes the relevant consonant will favour the embracing of innovative katakana that represent absent sequences. Such forms are expected to be adapted into Japanese earlier than the novel sequences which have no native sequence in any vowel context. For example, the existence of native sequences /fu/, /ta/, /te/, and /to/ have influenced the early integration of innovative /fa/, /fi/, /fo/, and /ti/ into Japanese as opposed to novel /v/-vowel sequences where there is no sequence of /v/ followed by a vowel which occurs in native vocabulary items. As shown in Table 2.5, /va/, /vi/, /vu/, /ve/, and /vo/ sequences are integrated later, in the 1960's, due to the

phenomenon due to the interchangeable sounds of /fu/ and /hu/ for the same character of /fu/ in Table 2.4.

absence of native /v/ in the Japanese system.

Table 2.5: The corresponding native sequences for some innovative katakana imported in the 1950's and the 1960's (*Source: Modified from Stanlaw's (2002, p. 569) innovative katakana chart)*

1950's	Native sequence(s)	1960's	Native sequence(s)
/ʤe/	/ʤa/, /ʤu/, /ʤo/	/ʃe/	/ʃa/, /ʃi/, /ʃo/, /ʃu/
/ʧe/	/ʧa/, /ʧu/, /ʧo/	/va/, /vi/, /vu/, /ve/, /vo/	-
/ti/	/ta/, /to/, /te/		
/di/	/da/, /do/		
/fa/, /fi/, /fo/	/fu/		
/wi/, /we/	/wo/, /wu/		

2.4.4 Focus of this Study: /w/ and /v/ variables

The adaptations of /fi/ and /fo/ in the 1950's (Stanlaw, 2002) show that Japanese has a higher tendency to select innovative forms for loanwords containing /f/-vowel sequences. These findings regarding /f/ sequences led me to examine glide sequences because /w/ sequences, like /f/ sequences, have some embedded sequences in the native inventory. In addition to examining the rate of use of innovative katakana in /w/ sequences, I will also compare the orthographic realization of /w/ sequences from loanwords with the realization of novel /v/ sequences. Recall that, unlike /w/, /v/ does not occur in any sequence in the native phonological system.

Table 2.6 illustrates that novel sequences including the /w/ glide in modern Japanese also show similar patterns to the innovative /f/ forms in terms of the introduction of innovative katakana and the patterning of /w/ in the native phonology. Because of the existence of three monomoraic sequences including /w/ in the native phonological system (/wa/, /wu/, and /wo/), I expect that innovative monomoraics such as /wi/ and /we/ should be preferred to conservative bimoraic realizations in the loanword system since the native phonology already includes some /w/-vowel sequences. The bimoraics /ua/, /ui/, /uu/, /ue/, and /uo/ are categorized as conservative forms as they are created by incorporating the native /u/ with all vowels.

Table 2.6: Conservative (bimoraic) /u/ and innovative (monomoraic) /w/ katakana variants

	Bimoraic (/u/+vowel)		**Monomoraic (/u/+ reduced vowel= /wV/)**	
Vowel		**katakana**		**katakana**
ア/a/	ua	ウア	**wa**	**ワ or ウァ**
イ/i/	ui	ウイ	wi	ウィ
ウ/u/	uu	ウー	**wu**	**ウ**
エ/e/	ue	ウエ	we	ウェ
オ/o/	uo	ウオ	**wo**	ウォ or **ヲ**

It should be noted that in Japanese phonology, monomoraic /wo/ is a native sound sequence with multiple possible orthographic representations as shown in Table 2.6. The katakana character of ヲ/wo/ is not used in any loanwords because it represents a particular morpheme, a case particle, and is always written using the hiragana character, を in Japanese. This particle を has another possible sound of /o/ [10] which is used more frequently by modern Japanese speakers. Since the native hiragana character of を/wo/ or /o/ is only used to represent Japanese grammatical functions, the katakana counterpart (ヲ) of hiragana を/wo/, is rarely used in modern Japanese. The character of ウォ/wo/ in Table 2.6, therefore, is marked as an innovative variant which is used to replace foreign /wo/ as in the word of "water"[11] in (2.11).

	Donor word	**Conservative**	**Innovative**
(2.11)	water	**ウオ**ーター/uo:ta:/	**ウォ**ーター/wo:ta:/
		*__ヲ__ーター/wo:ta:/	

Monomoraic /wa/ has two different characters in Japanese orthography: the right ワ/wa/ in Table 2.6 is the native character, and the left ウァ /wa/ is the innovative character. In the loanwords

[10] Note that the vowel /o/ (hiragana character of お or katakana character of オ) is also pronounced as /o/.

[11] There are a few possible vowels for the sequence of /wɔ/ in "water" depending on the dialect word has been adopted from the conventionalized Japanese borrowing.

that are adapted with a /wa/ sequence, the Japanese favour the traditional character of ワ/wa/[12]. In addition, the traditional Japanese system has another character of は /ha/ which is also realized as /wa/ when it functions as a particle that indicates the main topic in a phrase. However, the counterpart character in katakana form, ハ /ha/ is never adapted as /wa/ in loanwords. It remains as /ha/ all the time.

In addition to the innovative characters used in the writing of /f/ sequences (as in Table 2.4) and /w/ sequences (as in Table 2.6), the labiodental fricative consonant /v/ is also written as a set of innovative katakana characters which are created to represent the /v/ sound found in English and other foreign languages. Due to the absence of /v/ in the native system, the novel characters undergo two levels of innovation to mirror the foreign /v/. First, the core character of /vu/ for other /v/ variants is created from the affiliation of the symbol /u/ and a voicing diacritic (゛).

(2.12) Creation of core sequence of /vu/:
Native ウ/u/ + a voicing diacritic (゛) → ヴ/vu/ (core sequence)

The second stage is the combination of /vu/ with other half-sized vowels, excluding the vowel /u/, to form innovative sounds (bimoraic or monomoraic) with /v/ as in (2.13).

(2.13) Core ヴ/vu/ + half-sized vowels ァ/a/, ィ/i/, ェ/e/, ォ/o/

→ ヴァ/va/, ヴィ/vi/, ヴェ/ve/, ヴォ/vo/

Loanword /v/ is typically adapted as /b/ in Japanese, giving /ba/, /bi/, /bu/, /be/, and /bo/ as the conservative counterparts for /v/. The inventory of innovative /v/ characters and their conservative /b/ counterparts are recorded in Table 2.7.

[12] The preliminary finding in the section 3.2.2 shows that the innovative variant, ウァ /wa/, is not used at all in some loanwords.

Table 2.7: Innovative katakana inventory for bimoraic and monomoraic /v/ forms and conservative counterparts /b/ forms

	/b/ variants		Bimoraic[13] (devoiced /u/+ vowel)		Monomoraic (devoiced /u/+ rubic vowel)	
Vowel		**katakana**		**katakana**		**katakana**
ア/a/	ba	バ	**vua**	**ヴア**	va	ヴァ
イ/i/	bi	ビ	**vui**	**ヴイ**	vi	ヴィ
ウ/u/	bu	ブ	**vuu**	**ヴー**	**vu**	**ヴ**
エ/e/	be	ベ	**vue**	**ヴエ**	ve	ヴェ
オ/o/	bo	ボ	**vuo**	**ヴオ**	vo	ヴォ

The lack of a labiodental fricative in the native phonemic inventory constrains the use of /v/ in loanwords. The bilabial plosive /b/, on the other hand, occurs in the native phonological system and it used as an adaptation of foreign /v/ in Japanese. Fais et al. (2005) argue that the use of /b/ as an adaptation for foreign /v/ is reliably perceived by Japanese speakers because they are alert to the absence of /v/ in the native phonology. Therefore, the presence or absence of native phones or characters in Japanese phonological and orthographic systems is an important reason that this study analyzes the /v/ and /w/ variables and variants specifically.

Even though the monomoraics of /va/, /vi/, /vu/, /ve/, and /vo/ have been imported as innovative katakana since the 1960's (Stanlaw, 2002), many innovative katakana studies do not show wide diffusion of /v/. For instance, Sakamoto (2002) assesses the distribution of integrated katakana /b/ and /v/ for the loanwords of "violin" and "Beethoven", and finds that the register of newspapers uses more conservative /b/ than innovative /v/ as compared to internet material. Irwin (2011) also quotes the 2002 survey results by NHK (Nippon Hoso Kyokai or Japan Broadcasting Corporation), which indicate that Japanese people have a higher tendency to preserve the traditional /b/ forms rather than the novel /v/ forms for the same loanwords: "violin" and

[13] Unlike the bimoraic forms in Tables 2.4 and 2.6, I consider these bimoraic /v/ forms as innovative forms because they are combined using the core form of innovative /vu/ and vowels. See more discussion in section 3.4.1.

"Beethoven".

The shaded bimoraic /v/ column in Table 2.7 lists the interchangeable counterparts for the monomoraic group. The ratio of innovative bimoraic or monomoraic /v/ will be examined in this study to assess the distribution of native /b/ variants and novel /v/ variants.[14]

2.4.5 Hypotheses

A review of the literature has led to the hypotheses to be considered in this study. First, based on the presence of /wa/, wu/, and /wo/ sequences in the native inventory, I hypothesize that the innovative monomoraic characters used for the glide /w/ are the dominant option in Japanese as opposed to their conservative bimoraic counterparts. The innovative /w/ variants are expected to pattern like the innovative /f/ forms studied by Stanlaw (2002), principally because of the homologous presence of phonetic [fu] in the native sequences. In this case, the native sounds of /wa/, /wu/, and /wo/ are assumed as the trigger for other innovative /w/ characters (/wi/, /we/, and /wo/) in contemporary Japanese, especially in loanword adaptation.

The second hypothesis is that conservative /b/ variants are the predominant choice in foreign /v/ adaptation as compared to the novel and more accurate /v/ variants. This hypothesis follows from the absence of /v/ in the native phonological system.

2.5 Social Influences in Innovative Katakana Reform

The role of social factors is significant in the process of reforming the katakana characters in contemporary Japanese writings. Those factors include the bibliographical background of the writers, genres of the publications, as well as the role of bureaucracies who recommend particular

[14] See the section of 3.4.1 (Table 3.3) for the unremarkable usage of bimoraic innovative /v/ variants in the corpus.

guidelines in the use of innovative katakana and conservative katakana. In order to examine these social factors, I will use the framework of variationist sociolinguistics.

2.5.1 Variationist Sociolinguistics

Variationist sociolinguistics examines the role of social and linguistic factors in shaping language variation and change. The social factors commonly considered in sociolinguistic studies include social class, sex (or gender), age, geography, ethnicity and culture, style and register, social network, and communities of practices (Trudgill, 2000; Tagliamonte, 2011; Gordon, 2013). The fundamentals of variationist sociolinguistics are based on the observations and interpretation of language change and variation. Sociolinguistic studies also investigate the role of linguistic factors and have shown that aspects of linguistic structure shape quantitative, variable, linguistics patterns as well as categorical ones.

Traditional variationist sociolinguistics emphasizes verbal and face-to-face communications in field research (e.g. Labov, 1984; Tagliamonte, 2011). However, sociolinguists are using written corpora more frequently in recent work. According to Baker (2010), corpus analysis is indeed helpful to variationist sociolinguistics because corpora can contribute a tremendous amount of "existing data, along with computational tools and procedures which allow common (and rare) language patterns and frequencies to be identified quickly and accurately and compared across different populations" (p. 9). This study, therefore, uses a written corpus as a source of data to evaluate the usage of innovative and conservative katakana forms across different social factors.

Most previous literature on Japanese orthography focuses on the contextual usage and function of katakana characters in Japanese, as opposed to variation in their written form. Katakana

words are well-known for filling lexical gaps, giving special effects (such as creating a western atmosphere and showing trendiness (Shaad, 1987) and modernity, etc.), euphemisms (Rebuck, 2002), decorating advertisements (Dougill, 2008), and as an English learning tool for Japanese learners (Daulton, 2008 & 2015; Barrs, 2011; Champ, 2014).

2.5.2 Gender

Gender is a significant factor in sociolinguistic studies. Sociolinguistic studies have argued that males and females use "different varieties" in their speech even though they speak the same language (Trudgill, 2000, p. 65).

For example, Labov's (2001) work on the alternation of suffix *-ing* between men and women shows that males in Philadelphia use less of the prestige variant [ɪŋ] than females. Instead, men prefer the stigmatized form of [ɪn] when saying the *-ing* suffix. With respect to the current study, these works suggest that female and male writers may be expected to differ in terms of their preference for one katakana variant over another. As generalized by Tagliamonte (2011): "women tend to avoid stigmatized forms" (p. 32). In this study, I therefore expect female writers to choose the conservative katakana forms at higher rates than male writers. This is because conservative katakana characters are the prestige forms in the Japanese native system as Japanese bureaucracies who published katakana guidelines (Monbusho, 1991; Bunkacho, 1991; NINJAL, 2003a & 2003b & 2004 & 2006; JTCA, 2015) encourage the Japanese people to use conservative katakana forms as far as possible as opposed to innovative forms.

Although women tend to use prestige forms at higher rates than men, the effect of gender on variable use is complicated by the fact that women also tend to lead linguistic change (Labov 2001). Women may thus be expected to use more innovative forms if the innovation is a change in

progress. I interpret the innovative forms as stigmatized variants, in accordance with the published katakana guidelines discussed above, and expect women to use these innovative forms at lower rates than men. However, some time periods included in the corpus may represent a change in progress for the use of innovative katakana, complicating the predictions with respect to gender.

2.5.3 Age

Linguistic variation can be correlated with a speaker's or writer's age (Tagliamonte, 2011). In order to look for the varieties used by different generations, the patterning of linguistic features is usually compared with reference to the apparent time hypothesis. According to the apparent time hypothesis, the speech of older speakers represents the speech of earlier periods in time, equivalent to the state of the language when speakers were acquiring it in their youth. Various sources suggest that age plays a significant role in patterns of orthographic language use among Japanese writers. For example, results from BKK's Opinion Poll in 2012 show that more than 80% of people over age 50 do not understand the meaning of katakana words. This means that older generations face great difficulty when reading or listening to gaikokugo (katakanized foreign words) and new katakana words as they have minimal or no English proficiency as compared to young people.

Young speakers (Rebuck, 2002; Dougill, 2008; Barrs, 2011; Irwin, 2011; BKK, 2012; Inagawa, 2015; Kubozono, 2015a & 2015b) are treated as the salient group who are integrating novel sounds and sound sequences into contemporary Japanese without modification. In other words, young Japanese speakers and writers are more likely to violate Japanese phonotactics in order to more closely approximate loanword forms as pronounced in the donor language. However, as stated by Matsuda (2003), young Japanese would learn to use prestige forms in their speech under “strong peer pressure” when “they are introduced to wider social classes in later adolescent

years" (p. 131). This suggests that age-grading is possible. Age-grading is when speakers adjust their linguistic output as they age (Tagliamonte, 2011; Labov, 1994). But, this study focuses mainly on the examination according to the apparent time hypothesis. Therefore, the apparent time hypothesis is the young Japanese writers are expected to embrace and use more innovative variants than the old writers.

In addition to looking at speaker age, use of corpus data allows me to consider real-time data by comparing the use of innovative katakana across texts published in different years. The time of publication could potentially reflect the diachronic and synchronic variation of the use of innovative variants. As compared to apparent time data that can be garnered by looking at writers of different ages, access to data published at different times provides real-time data that can provide evidence regarding changes in use of katakana over time. This study predicts that innovative variants are more likely to appear in recent publications.

2.5.4 Register

Register refers to the genres of publications, specifically the written genres for both traditional and web-based publications. Depending on the nature of the publication and the targeted audience or readers, the language use is expected to vary across different registers. This is likely because "people tend to use higher prestige variants more often in more formal styles and lower prestige variants more often in informal styles" (Tagliamonte, 2011, p. 34). A previous study by Sakamoto (2002) on Japanese publications have provided evidences to support Tagliamonte's (2011) statement.

Sakamoto (2002) examined the distribution of conservative /b/ and innovative /v/ in three daily Japanese national newspapers and three web search engines. His findings clearly show that

all three newspapers prefer the conservative /b/ forms /baiorin/ and /be:to:ben/ when foreign words of “violin” and “Beethoven” are borrowed into Japanese. Only about 10% of the loanwords use the innovative forms of /vaiorin/ and /be:to:ven/. He further compared the /b/ and /v/ distribution patterns to the results from three online searching tools, i.e., *Yahoo!*, *goo*[15] and *Google*. The result from web searching engines shows that innovative forms of /v/ loanwords comprise almost 50% of the search results. For proper nouns like 'Beethoven', the innovative /be:to:ven/ is preferred over /be:to:ben/.

Sakamoto's (2002) results indicate that the writers or editors of newspaper companies (traditional and formal register) are expected to use conservative and standard language whereas the scribes of the internet (modern and informal register) could have a lesser concern with the conservation of standard Japanese. Perhaps, another possible reason for the substantial innovative usage on the internet is the writers are comprised of not only native Japanese but also non-native writers who have a high proficiency in Japanese. Moreover, internet writers are very likely exposed to more English from native English writers than the newspaper's editors since the internet was first introduced by the west and utilized English as the predominant language.

2.5.5 Bureaucracies

Since 1955, Japanese bureaucracies started publishing recommendations for katakana orthography (Irwin, 2011) with the aim to standardize katakana forms in the Japanese language. Japanese bureaucracies such as Monbusho (1991), Bunkacho (1991), NINJAL (2003a & 2003b & 2004 & 2006) and JTCA (2015) encourage Japanese writers to employ the conservative katakana forms whenever possible in order to preserve the standardization of conservative variants used in

[15] *goo* is a Japanese online search engine powered by *Google*. This internet search engine is used to crawl and index Japanese websites.

Japanese. Therefore, it is not surprising to note Sakamoto's (2002) finding that showed national newspapers publishers prefer the conservative /b/ forms over the innovative /v/ forms for "violin" and "Beethoven". However, recently, the NHK (2015) Broadcasting Terminology Committee has suggested that innovative katakana forms should be considered for recently imported words due to the phonological and orthographic similarities between donor sounds and novel forms. For example, the innovative form of /webu/ has the closest sound and romaji orthography with the English "web" as compared to the conservative form of /uebu/. On this basis, NHK (2015) encourages Japanese writers and speakers to opt for the innovative variants of /webu/ instead of the conservative counterpart of /uebu/.

2.5.6 Hypotheses

With respect to gender, male writers are expected to use more innovative katakana variants than females, based on literature from Trudgill (2000), Labov (2001), and Tagliamonte (2011). Regarding the findings from previous literature for age (Rebuck, 2002; Dougill, 2008; Barrs, 2011; Irwin, 2011; BKK, 2012; Inagawa, 2015; Kubozono, 2015a & 2015b), I expect young writers, as opposed to older writers, to produce more novel katakana characters.

Another hypothesis predicts that the traditional and formal registers of the corpus sources will use more conservative katakana forms relative to the modern and informal archives (Sakamoto, 2002; Tagliamonte, 2011). Finally, the distribution of novel forms (monomoraics of /w/ and /v/ katakana) is expected to increase diachronically based on the recent recommendation by NHK (2015).

CHAPTER 3: METHODOLOGY

This chapter consists of four sections: The first section explains the features of the corpus used in this study. Section 3.2 is the description of the research focus, including linguistic and social factors. Section 3.3 details the procedures used in collecting and coding the data I extracted from the corpus. In the last section, I describe the methodology of the statistical analysis and the limitations of the study.

3.1 The Corpus: BCCWJ

I obtained the data from the Chunagon corpus which is one of the interfaces in The Balanced Corpus of Contemporary Written Japanese (BCCWJ). The BCCWJ corpus is one of the corpora compiled by the National Institution of Japanese Language and Linguistics (NINJAL) since 2006.

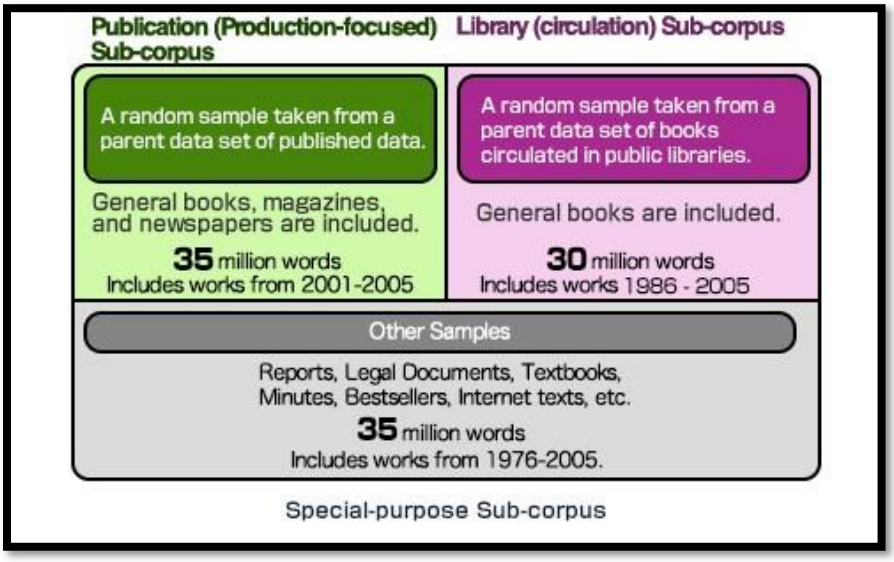

Figure 3.1: The registers and the corresponding timelines of sampled data in BCCWJ *(sourced from Basic Design Policy Balanced Corpus of Contemporary Written Japanese- http://pj.ninjal.ac.jp/corpus_center/bccwj/en/basic-design.html)*

The size of the BCCWJ is 104.3 million words which are extracted from a wide range of

published genres as listed in Figure 3.1[16]. NINJAL has been using a random sampling method to excerpt samples of each genre in order to compile the corpora. Random samples are chosen from various sources as "Fixed-Length" samples (short samples of 1,000 characters for each) and "Variable Length" samples with a length limit of 10,000 characters. "Fixed-Length" samples have consistency in size and are appropriate for statistical analyses of variation whereas "Variable Length" samples are appropriate for discourse analyses. The Chunagon corpus is compiled using "Fixed-Length" samples.

The parent data sets in BCCWJ consist primarily of printed forms of publications which are written for a large number of readers. For instance, books, newspapers, magazines, business reports, blogs, internet forums[17], and legal documents are included in the corpus. The parent sets in the database exclude some non-public or unofficial sources such as letters, diaries, billboards, packages, private journals, and messages, all of which contain personal writing styles. Figure 3.1 also shows that the overall sampled materials are selected from the 1970's to 2000's with the earliest year of 1976 and the latest year of 2005[18].

The corpus also includes detailed bibliographical information such as registers, authors' gender and age (interpreted from Year of Birth or YOB), genres, publishers, and publication year as illustrated in Figure 3.2. These bibliographical data that come along with the morphological information make this data source ideally compatible with sociolinguistic research.

[16] Random samples in BCCWJ are divided into three sub-corpora as in Figure 3.1. But in order to accommodate the focuses of this research, this study has different categorizations as described in section 3.2.3.
[17] Although blogs and internet forums are not physical printed forms, these types of publications are considered as written materials as well. Further descriptions of the corpus's "Data Sampling" are available on its website: http://pj.ninjal.ac.jp/corpus_center/bccwj/en/sampling.html
[18] The earliest and latest years are inaccurate because the searched results from the corpus have shown 1975 was the earliest year while 2008 was the latest year. But, it should be noted that there is no data for the years of 2006 and 2007.

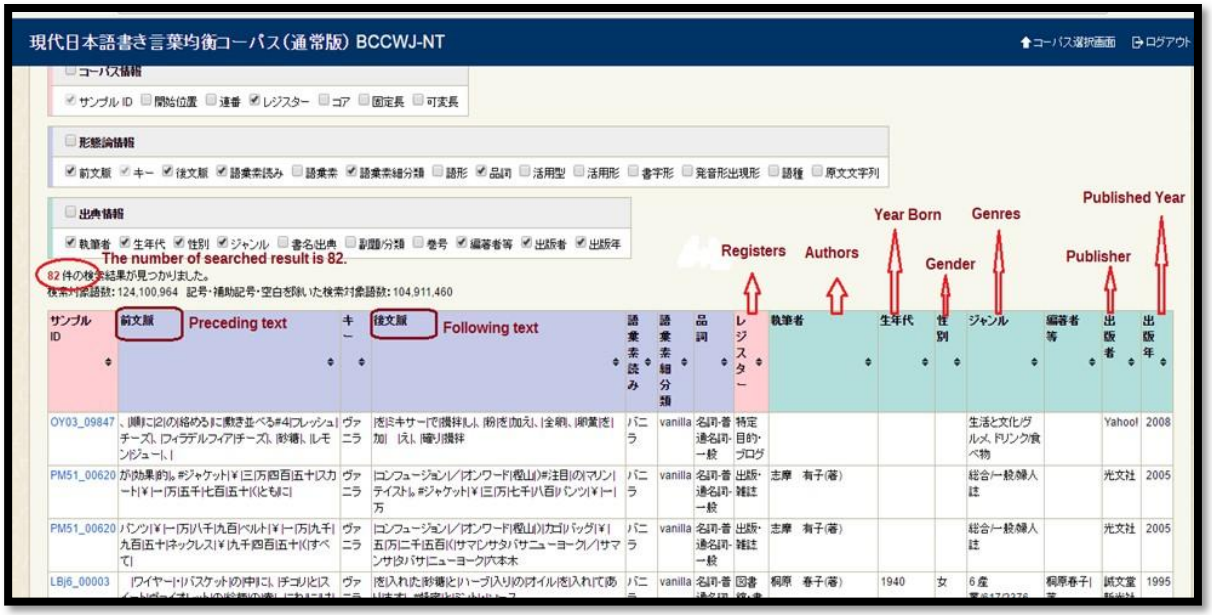

Figure 3.2: Bibliographical information in the search result of innovative /vanira/ "vanilla"

3.2 Focuses of Analysis

As discussed in the preceding chapter, this study examines orthographic variants of /w/ and /v/ considering not only the linguistic factors but also the social factors as described in the following sections.

3.2.1 The Phonological (and Orthographic) Variables

In this work, I focus only on /w/ and /v/ variables from English donor words which have conservative and innovative variants in Japanese as shown in Table 3.1. The orthographic characters for innovative variants of /w/ and /v/ variables are derived from the native character of the vowel ウ/u/, except for the character of ワ/wa/, and the character of ウ/wu/ which shares the same character as the vowel /u/.

Table 3.1: The conservative and innovative variants for /w/ and /v/ variants in Japanese

Variable	Variants in Japanese Katakana			
	Conservative		Innovative	
/w/	/u/ variants		/v/ variants	
	/ua/	ウア	/wa/	ワ or ウァ[19]
	/ui/	ウイ	/wi/	ウィ
	/uu/	ウー	/wu/	ウ[20]
	/ue/	ウエ	/we/	ウェ
	/uo/	ウオ	/wo/	ウォ
/v/	/b/ variants		/v/ variants	
	/ba/	バ	/va/	ヴァ
	/bi/	ビ	/vi/	ヴィ
	/bu/	ブ	/vu/	ヴ[21]
	/be/	ベ	/ve/	ヴェ
	/bo/	ボ	/vo/	ヴォ

3.2.2 Linguistics Factors

As mentioned above, the variants of /wa/ and /wu/ are present in the native inventory with the conservative form of /wa/ and the vowel /u/. The phonetics of katakana characters ワ/wa/ and ウ /wu/ (the same character as vowel ウ/u/) are considered present in the native Japanese system. There is another variant of /wo/ that is marked as a native sound even though the native /wo/ is represented by a different katakana character, ヲ[22]. However, there are no /wi/ and /we/ variants or /v/ variants in the conservative phonological and orthographic system.

One hypothesis examined in this thesis deals with the presence of these forms in the Japanese phonological system, which predicts /wa/, /wu/, and /wo/, would increase the probability of using innovative /w/ variants as compared to the conservative variants of /u/. Another expectation is that the lack of labiodental sounds and characters for /v/ variants in the native

[19] The predicted innovative ウァ/wa/ variant and the native ワ/wa/ are excluded in the research due to the reason stated in section 3.2.2.
[20] The /wu/ variant shares the same katakana character with vowel /u/ and/wu/ is not included in this paper as well (see section 3.2.2).
[21] Please refer to the 2-level emergence of /v/ variants in section 2.4.4: (2.12) and (2.13).
[22] Please refer to section 2.4.4 for the explanation of /wo/ functions in Japanese.

systems leads Japanese speakers and writers to avoid the innovative /v/ variants and favour the traditional /b/ variants.

For example, this study predicts a lower likelihood of using the conservative /uo:ta:/ for "water", as compared to the novel /wo:ta:/, due to the presence of /wo/ in the native phonological system. And the probability of writing the conservative /banira/ is predicted to be higher than the innovative /vanira/ for "vanilla"' as /v/ is absent in the traditional Japanese phonology.

This study concentrates only on /wi/, /we/, and /wo/ variants because the variant of /wu/ is not an innovative form and its katakana character is identical to the katakana character for the vowel /u/. Unlike the /wu/ variant, which utilize the character of the vowel /u/, there are possible native /wa/ and innovative /wa/ variants which can be used in loanword adaptation as shown in Table 2.6. Therefore, to validate the exclusive use of traditional variant of ワ/wa/, this study has preliminarily tested the usage of innovative /wa/ with a loanword that has a /wa/ variant. A few loanwords with the /wa/ segment, as shown in (3.1), (3.2), and (3.3), were tested in the database. As expected, the database showed only the results with traditional ワ/wa/ variants and zero results for innovative ウァ/wa/ variants. This preliminary finding has led me to rule out the variant of /wa/ in this study as well.

Loanword	**Conservative**	**Innovative**
(3.1) wonderful /wʌndərfəl/	**ワ**ンダフル /wandafuru/	*__ウァ__ンダフル /wandafuru/
(3.2) one piece /wʌn pis/	**ワ**ンピース /wanpi:su/	*__ウァ__ンピース /wanpi:su/
(3.3) one pattern /wʌn pætərn/	**ワ**ンパターン /wanpata:n/	*__ウァ__ンパターン /wanpata:n/

The second linguistic factor considered in the study is the position of /w/ and /v/ variables

in English loanwords. I selected loanwords which have /w/ and /v/ variables in the different positions such as word-initial, word-medial, and word-final as shown in (3.4) and (3.5). However, not all /w/ variants occur in all positions. For example, /w/ variants in (3.4) occur only in initial and medial positions but never in final position.

	Initial	**Medial**	**Final**
(3.4) /w/:	wine	sweater; power	(not applicable)
(3.5) /v/:	volunteer	motivation	love

To my knowledge, no serious attempt has been made in assessing the effect of position within a word in loanword adaptation with respect to Japanese orthography. Therefore, I referred to Lively et al.'s (1994) work on training Japanese speakers to distinguish English /l/ and /r/. Their findings show that word-final /l/ and /r/ have the highest probability of correctness, followed by word-initial /l/ and /r/, and lastly /l/ and /r/ in word-medial position. Based on this observation, I expect the likelihood of the initial /w/ variable to be adapted as the innovative variant to be higher than that of the medial /w/ variable. For /v/, which could occur in all positions in English as /l/ and /r/ do, I predict the probability of using innovative /v/ variants should be consistent with the correctness of /l/ and /r/. This means that the preference of using innovative /v/ in final position should be the greatest, followed by initial /v/ innovative usage, and then the use of medial /v/ innovative forms.

3.2.3 Social Factors

Apart from determining the linguistic factors that influence the patterns of innovative katakana forms, I also study the effect of several social factors on the usage of innovative katakana sequences.

Demographic variation

First, I assess the trend of conservative and innovative sequences based on the writers' demographic background, namely gender and age. I pay attention to the search results with single author and multiple-authors who are all in the same gender group. There are 3 groups of gender classification: Male (M), Female (F), and Not Applicable/ Unknown (N). N refers to the multi-authors that are female(s) and male(s), as well as the writers whose gender category is not available in the corpus. Based on the argument by Trudgill (2000) as noted in section 2.4, women writers are expected to use more conservative forms than men writers.

Concerning the second demographic factor, age, this study analyzes single writers' year of birth (YOB). This means that age was not considered for work produced by unknown or multiple authors as there is no one YOB. Also, some very early YOBs (1860, 1870, 1880, and 1890) were combined into a group of "1800s" because a smaller group of authors were present in data from those decades. Therefore, this study examines the age variation in using innovative variants by considering the YOB for single authors in 11 decades (as in Figures 4.4 and 4.12), starting from the 1800's, continued with the 1900's until the 1990's. According to the statistical data from BKK's Opinion Poll (2012), I expect to see younger authors use more innovative variants than the older writers.

Variation across registers

This study also investigates the probability of innovative and conservative katakana forms across different genres. The genres provided by the database include many categories, some of which share similar criteria. For this reason, I rearranged the 13 types of genres given in the database into 4 main registers as shown in Table 3.2. The genre of "verse" is eliminated from the data due to the extremely low frequency in the search results.

Table 3.2: Four types of register groups and the corresponding genres

	Register	Genres	Timeline
1	Book	Publication Books; Library Books; Bestseller Books; Textbooks	1970s to 2000s
2	Publication	Newspapers; Magazines; Public Papers	2000s
3	Web	Yahoo! Blogs; Yahoo! Answer	2000s
4	Other	Whitepapers, Legal Documents; National Assembly Proceedings	1970s to 2000s

Regarding Sakamoto's result (2002), I believe that informal registers such as Web (*Yahoo! Answer* and *Yahoo! Blogs*) will include more innovative forms. Tagliamonte (2011) also asserts that "people tend to use higher prestige variants more often in more formal styles and lower prestige variants more often in informal styles" (p. 34). In this paper, the prestige forms are the conservative katakana characters that are expected to appear more in Book and Other as well as in Publication.

Diachronic variation

I also assess diachronic variation using the information of "Published Year" (henceforth Publication Year). The expectation is recent publications should have a higher preference for the use of innovative forms than older publications.

3.3 Data Collection

I registered with NINJAL in order to access the corpus online and downloaded all relevant search results. I selected a total of 60 katakana words, provided in Appendix 1, from previous literature (NINJAL, 2003a & 2003b & 2004 & 2006; NHK, 2006-2017; Daulton, 2008; Irwin, 2011; JCTA, 2015). These include 44 katakana words which contain /v/ sequences and 16 which contain /w/ sequences. These loanwords were chosen based on the following criteria.

First, the katakana forms must be unambiguous with respect to the source word in the donor language. Second is whether the words exhibited variation (conservative and innovative).

	Donor word	Conservative	Innovative
(3.6)	window	ウインドウ/uindou/ ウインドー/uindo:/	ウィンドウ/windou/ ウィンドー/windo:/
(3.7)	*warming up	＊ウオーミンアップ */uo:minappu/	ウォーミンアップ */wo:minappu/
(3.8)	*van	バン/ban/	ヴァン/van/

For instance, “window” in (3.6) is a valid entry for this study because it has both katakana variants. But, “warming up” is excluded in data collecting due to the zero result for all possible katakana transliterations as shown in (3.7). Donor words like (3.8) were not extracted as data due to the ambiguity of the conservative katakana characters. Katakana renderings such as /ban/ actually can be reverse-transliterated into another English word: “bun”, in addition to “van”.

As mentioned above, for /w/-based loanwords, 16 of 44 loanwords are considered primary while another 18 are derivations and homophones of the primary 16. The primary words are found in Appendix 1 and the derived words and homophones are found in Appendix 2. For example, searching the database for “week” not only shows results of “weekend”, “weekender”, “weekday”, and “weekly but also its homophone “weak” and “weakness”, a form derived from “weak”. Another example is “wait” which is the only primary donor word that has no results from the corpus; instead the searched tokens are from its derived word of “waitress” and homophone of “weight”. This study skipped innovative /wa/ and /wu/ variants because foreign /wa/ and /wu/ sequences are always adopted as only traditional variants of ワ/wa/ and ウ/u/ respectively (see section 3.2.2).

Similar to what was described above for /w/, 44 of the 88 /v/-based loanwords are considered primary while another 44 are derived from the primary words. For example, searching the database for “view” also returns results of “viewer”, “review”, “interview”, and “interviewer”.

	Donor word	Conservative	Innovative
(3.9)	service	サービス/sa:bisu/	サーヴィス/sa:visu/
(3.10)	motivation	モチベーション /moʧibe:ʃon/	モチヴェーション /moʧive:ʃon/
		モティベーション /motibe:ʃon/	*モティヴェーション */motive:ʃon/

As shown in examples (3.6), (3.9), and (3.10), all these loanwords contained both innovative forms and conservative forms. Some loanwords such as “window” and “motivation” have more than one conservative variant due to the variation occurring in other segments of those words. For example, “window” in (3.6) has another set of variants with long vowel in final segment of /do:/ besides the final diphthong segment /dou/. And for “motivation”, the extra variation occurred on /ti/ (/ʧi/) segment.

Usually, if a loanword has multiple conservative variants, it will also have the same number of innovative counterparts. But, in this corpus, “motivation” shows only one innovative variant because there was no result for innovative form of /motive:ʃon/. Even so, these kinds of loanwords were included in this study as they show variation in the use of conservative and innovative forms. I categorized these additional katakana forms based on the variants of the segments focused in this study: /u/ vs. /w/ segments and /b/ vs. /v/ segments as shown in (3.6) and (3.10).

3.3.1 Searching Method

The BCCWJ corpus is designed based on morphological searching features. As shown in Figure 3.3, there were four types of searching methods. The “Short Unit Search” is used to examine individual units, specifically the smallest lexical items which convey meaning. The method of “Long Unit Search” is utilized to evaluate linguistic properties for the longer texts. “Position Search” uses properties of samples, i.e., Sample ID and Initial Position of the string, to look for

necessary results.

Figure 3.3: Four options to choose from search methods in the corpus

This study used the "Character String Search" method to carry out the data collection. This search method was the ideal option to obtain data because this study needed only the frequencies of the relevant loanwords' appearance in the corpus as well as the authors' demographic and publication information shown in the results.

Katakana words that are borrowed or derived from English have a distinctive linguistic property relative to native words in Japanese, namely marking of the unit boundaries. For example, in native Japanese words such as 愛煙家/aienka/ "heavy smoker", the boundaries between morphemes are clear as each Chinese character represents a meaning: 愛/ai/ is "love", 煙/en/ is "smoke", and the meaning of 家 /ka/ is "person(s)". However, since there is no morphologically significant "space" between units or characters in Japanese orthography, Japanese sometimes uses the symbol " ・ " to distinguish morpheme boundaries in katakana words as illustrated in "heavy smoker":

(3.11) ヘビー・スモーカー versus ヘビースモーカー
/hebi: ・sumo:ka:/ /hebi:sumo:ka:/

Accordingly, some inaccurate matches shown in the search results are due to the issue of the unit boundary. For instance, when searching for "wait" or "weight" (both loanwords share the same variants of /ueito/ and /weito/ in Japanese), I found that the conservative variant, /ueito/, in a racing horse name, コ**ウエイト**ライ/ko**ueito**rai/ also appeared in the search results. Another example is /ka:bu/, the conservative variant for "curve" which appeared in the phrase サッ**カーブ**ログ/sak**ka:bu**rogu/ ("soccer blog"). Of course, these incorrect results were excluded from the study.

However, I did include homophones and derivations from primary loanwords in this study. For example, one of the selected primary loanwords, "view" returned results not only for "view" but also its derivations, such as "review", "interview", and so on. The example for homophones could be found in "wait" and "weight" as they share the same katakana forms in Japanese.

3.3.2 Results Display Format

Figure 3.4 shows the location to insert the katakana words as well as the options for "Result display unit". I selected both options, i.e., short unit data and long unit data for searching. This indicates that I searched for each katakana word twice, once with the short unit option and again with the long unit. This is because this study needed the information from the "語彙素細分類" or "Lexeme Subclassification" column which is included in short-unit search results. This column provides equivalent English words (Figure 3.6) for katakana words in the "Keyword" column in short-unit results. Also, this study required the "Keywords/ Keyphrase" column in the long-unit results to

clarify the "correct"[23] katakana words in searched concordances.

Figure 3.4: Search box for katakana words and the options for Result display unit

Figure 3.5: Descriptions of the searching procedures

For the "Key display format" selection in Figure 3.5, I set the blue highlighted option, "Keywords", which displayed only keywords in the search results instead of the entire search string

[23] This is because the "Keywords/phrase" column in the long-units search results shows the whole word or phrase which is associated with the katakana word entered in the search box. This allowed me to eliminate irrelevant results, especially when the searched word is only a part of the word or phrase in the search results. For example, conservative /ueito/ for "wait" or "weight" which appeared in a horse name/ko**ueito**rai/ and /ka:bu/ for "curve" which also found in "soccer blog" /sak**ka:bu**rogu/ as mentioned earlier.

(the second option in the drop-down). I checked all the categories and subcategories of genres and registers under the section of "Search Targets". While in the "Search Operation" section, I changed the default number of 20 to the minimum number of words, 10, so that the length of preceding and following texts shown was restricted to 10 words.

After all search components were set, I started entering the katakana words and using the "Search" and "Searching Results" download buttons as shown in Figure 3.5 to acquire the search results.

3.3.3 Getting the Searched Results

As mentioned above, I downloaded the searched results from the database in order to extract and recode the relevant information for this study. The example shown in Figure 3.6 is the results for the conservative variant for "service", /sa:bisu/, which shows the highest frequency of search results (16425 cases) relative to other variants for other loanwords.

The original search results from the database were solely in Japanese. I eliminated the unnecessary columns and retained 6 relevant columns: "Sample ID", "Lexeme Subclassification", "Gender", "YOB", "Register", and "Published Year". I recoded "Register" into 4 main types of registers as discussed in section 3.2.2.

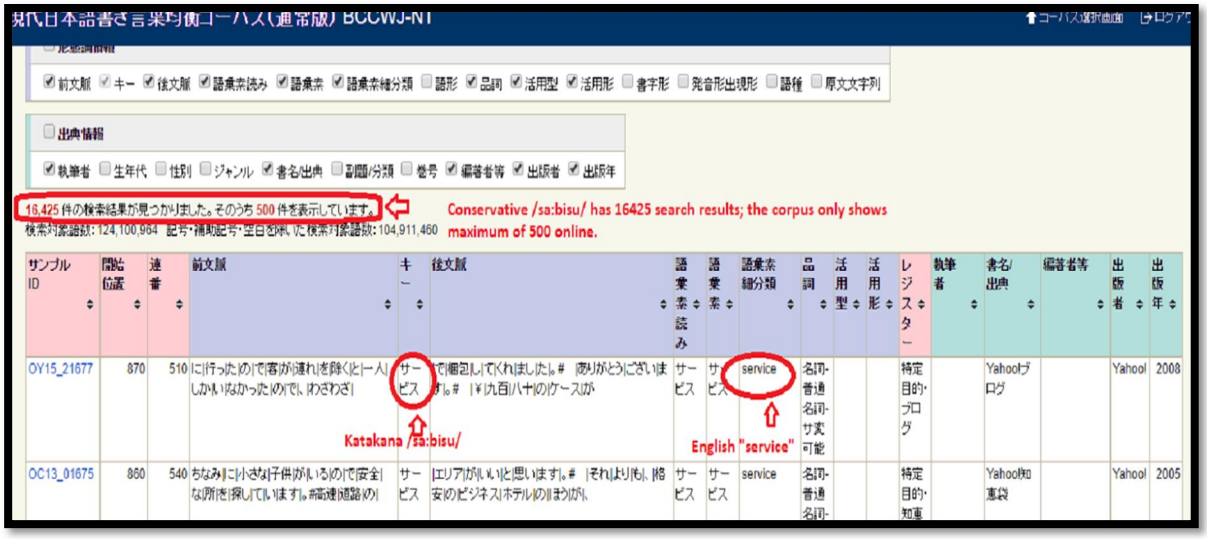

Figure 3.6: The search results of conservative katakana for "service"

Furthermore, I created five additional columns for "Katakana" (katakana in alphabetical symbols), "English Segment", "Segment" (referring to katakana segments), "Form" (either conservative or innovative), and "Position" for specifying the positions of /w/ and /v/ variants as initial, medial, or final.

3.4 Data Analysis

I used R (R Core Team, 2017) to implement the statistical analyses and to create figures visualizing the distribution of conservative and innovative forms according to linguistic and sociolinguistic variables. R was also used to realize a multivariate statistical analysis as a means to determine the influential factors for choosing innovative variants.

3.4.1 Limitations of the Database: Bimoraic Innovative /v/ Variants

As discussed in the previous chapter, in line with the conservative (bimoraic) forms of /ua/, /ue/, /ui/, and /uo/, it should be possible to constitute similar bimoraic forms for the /v/ variable: /vua/, /vue/, /vui/, and /vuo/ (Table 2.7). However, the preliminary search results for "Valentine" (3.10)

showed that, the monomoraic /v/-single vowel variants (/va/, /vi/, /ve/, /vo/) are more predominant than these bimoraic /v/-diphthong variants (/vua/, /vue/, /vui/, /vuo/).

	Donor word	Conservative	Innovative
(3.10)	Valentine	/barentain/	/varentain/ or */vuarentain/

I also tested the usage of bimoraic /v/ variants by only inserting the characters of the bimoraic innovative variants (/vua/, /vue/, /vui/, /vuo/) as listed in the second column from left in Table 3.3 and found that the frequencies obtained from the database are very low with most of them found in "Unknown" words as recorded in the searched results. Most of the "Known" loanwords are proper nouns, particularly places and personal names, as in recorded in Table 3.3.

As we can see, the examples of "Known" loanwords are mostly adapted from non-English languages, and this provided another reason to eliminate these novel bimoraic /v/ variants from this study which focuses on English-based loanwords[24].

Table 3.3: The findings of bimoraic innovative /v/ variants from the corpus

Bimoraic /v/ variants		Frequency			
		OVERALL	Unknown	Known	Examples of Known loanwords
vua	ヴァ	20	18	2	variation
vui	ヴィ	49	44	5	Bouillabaisse; Eve; Wilhelm
vue	ヴェ	25	8	13	Veda; Soviet; Hervouet; Velazquez; Venezuela; Novelli; Vietnam
vuo	ヴォ	9	7	2	vaudeville; Vương (Vietnamese surname)

3.4.2 Limitations of the Data: Social Factors

One shortcoming of the dataset concerns the lack of information regarding age and gender for many tokens. As shown in Table 3.4, only 39.4% of the overall search results were used to assess the gendered usage of innovative variants. This is because for many tokens, the gender of the writer

[24] The English words that correspond to the selected loanwords are treated as contemporary English words. The historical origin of the chosen English words is not considered in this study.

is not recorded in the database or the text is listed as having multiple authors with different genders. These tokens are coded with "N" for unknown gender in Table 3.4.

Table 3.4: The frequencies distribution for gender

GENDER	FORM		TOTAL	
	Conservative	Innovative	Case	%
F	5042	860	5902	8.3
M	18672	3376	22048	31.1
N	**38184**	**4794**	**42978**	**60.6**
			70928	**100.0**

Similarly, for the age factor, this study had to exclude 71.9% of the data (Table 3.5) when examining the distributional patterns of innovative practice in the corpus. As in the case for gender, the age of the writer is not known in many cases. When relevant information is unavailable, year of birth is listed as "0" in Table 3.5, below.

This means that the descriptive analysis of "Age" and "Gender" is based on a subset of the data which might distort the findings. However, in the multivariate analyses, all data was included with categories of "N" and "0" being used as the categories of unknown gender and age values.

Table 3.5: The frequencies distribution for age (YOB)

YOB	**/w/ Variant**		**/v/ Variant**		**Total**	
	Conservative	**Innovative**	**Conservative**	**Innovative**	**Case**	**%**
0	2069 **(67%)**	4701 **(72%)**	42798 **(73%)**	1421**(56%)**	**50989**	**71.9**
1860	0	1	0	0	1	0.0
1870	1	0	1	1	3	0.0
1880	0	1	2	1	4	0.0
1890	9	0	17	0	26	0.0
1900	6	5	57	12	80	0.1
1910	26	10	119	23	178	0.3
1920	108	87	962	106	1263	1.7
1930	193	263	2831	283	3570	5.0
1940	267	400	4441	273	5381	7.6
1950	168	462	4298	261	5189	7.3
1960	182	424	2637	127	3370	4.8
1970	54	137	602	25	818	1.2
1980	1	4	39	0	44	0.1
1990	1	2	9	0	12	0.0
	3085	*6497*	*58813*	*2533*	**70928**	**100.0**

CHAPTER 4: RESULTS AND ANALYSIS

In this chapter, I report the descriptive statistical results for the /w/ and /v/ variables across the linguistic and social factors. I also describe the outcome of multivariate regression analyses.

The first section explains the reasons for presenting the statistical results separately for the /w/ and /v/ variables instead of an overall descriptive finding. Section 4.2 presents the frequency of innovative and conservative variants for the /w/ variable. Section 4.3 does the same for the /v/ variable. Both sections include the multivariate statistics after the descriptive results. The results of multivariate analyses present the correlations between linguistic and social factors in the use of innovative variants.

The final section of this chapter provides a summary of the main findings.

4.1 Overall Results and Distributions for /v/ and /w/ Variables

Before turning to the quantitative data that shows the pattern in the use of innovative orthography in borrowed English words, I present a note about the presentation of the results in this chapter. A decision was made to address each variable separately instead of as a unified phenomenon.

Table 4.1: The counts for both variants of /w/ and /v/ variables

Variable	Variants Type		Frequency
	conservative	innovative	
w	3085	6497	9582
v	58813	2533	**61346**
Total	**61898**	**9030**	**70928**

Looking at /v/ and /w/ variables on their own reveals clear differences in the behaviour of each variable. This may be because of the sheer quantity of /v/ tokens in the data set. As shown in Table 4.1, nearly 86% (61346 cases) of the data consisted of /v/ forms whereas the /w/ variable covered only 14% (9582 cases) of all data. In order to properly assess the two variables, and to

avoid the overwhelming majority of cases from masking the smaller set of data, I chose to treat each variable under separate statistical analysis.

Apart from this large imbalance, the phonological position of the /v/ and /w/ segments in the loanwords, one of the factors under investigation here, presented another bias for an overall statistical analysis that integrated the two variables. While the /v/ segments can occur in all positions (initial, medial, and final), the /w/ segments only appear in initial and medial positions. This difference would present difficulties in interpreting the results of a test that looked at phonological position of both variables under a single analysis.

In addition, preliminary examination of the data suggests very different effects of social factors for the /w/ and /v/ variables. Important patterns could thus be missed if both types of variables were considered together.

4.2 Analytical Results for /w/ variable

Let's turn now to the descriptive findings for linguistic factors included in this analysis. As described earlier in the methods chapter, I have treated the dependent variables as having two variants, either conservative or innovative orthographic variants. The independent variables included whether the phoneme is present or absent in the native Japanese phonology and the position within a word where the variants is found. These results are followed by the findings from the analysis of social factors, including the Gender, Age (YOB) of the writer, the Register or genre of writing the variant was found in, and the Publication Year of the written text.

4.2.1 Orthographic /w/ Variants in Japanese

As mentioned in the previous chapter, the chosen loanwords have at least one conservative form and also at least one innovative form (see section 3.3). However, some loanwords appear with additional variants not predicted in advance. The borrowed forms of "sandwich" and "whisky" show an additional variant besides the conservative /ui/ variant and innovative /wi/ variant. This is illustrated in Table 4.2.

Table 4.2: Orthographical Variants for "sandwich" and "whisky"

	English variables	**variant 1**	**variant 2**	**variant 3**
sandwich	wɪ	sando**ui**ʧi	sando**wi**ʧi	sando**i**ʧi
whisky	wɪ	**ui**suki:	**wi**suki:	**uwi**suki:

In Table 4.2, we see that "sandwich" appears with an /i/ variant and "whisky" is spelled with the form of /uwi/. I categorized /i/ and /uwi/ as conservative variants because of the long historical reason of established usage of /i/ (NHK, 2012) and /uwi/. This is supported by the data from this corpus showing that "sandwich" and "whisky" did appear in the register of Book as early as 1975 (Appendix 5[25]). Another reason is the bimoraic structure of /uwi/ which is constituted by two moras: /u/ and /wi/. Recall that, in this study, the bimoraic forms are considered as conservative variants and the monomoraic forms are considered as innovative variants.

[25]Appendix 5 records the top 10 oldest /w/ and /v/ loanwords found in the corpus.

4.2.2 The Presence of Native /w/ Sequences

As predicted, the innovative sequences /wi/, /we/, and /wo/ should occur at comparatively high rates because similar forms containing /w/ are already present in the Japanese inventory. This prediction is supported by the data: 67.8% of the forms were found to be innovative, compared to 32.2% of forms which contain conservative variants (Figure 4.1).

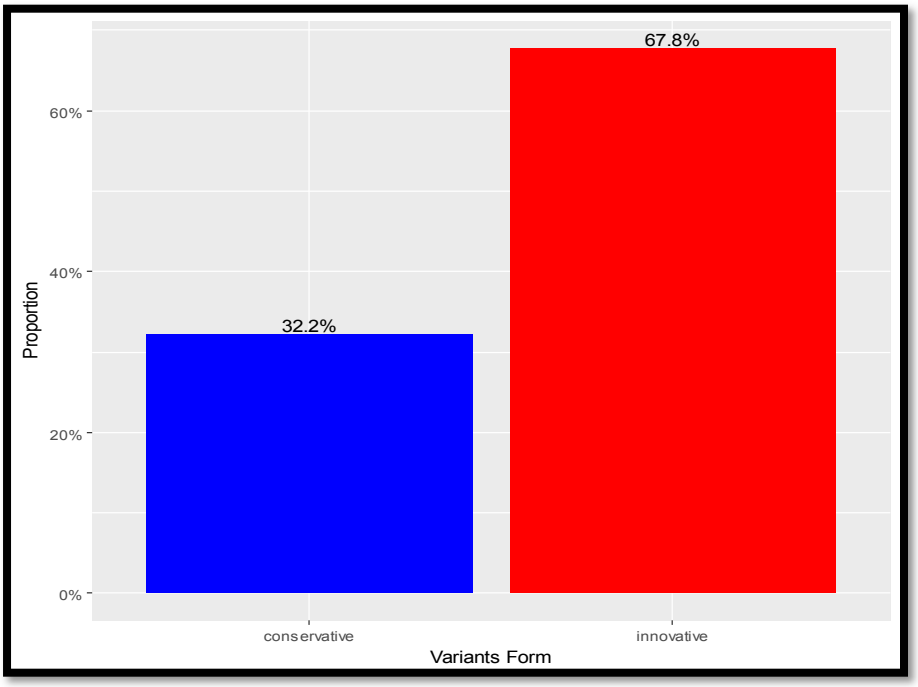

Figure 4.1: The overall distribution for /w/ variables

4.2.3 Positions of the /w/ Variables

Since the /w/ segment cannot occur in the final position, the results shown here are only for initial and medial positions. Figure 4.2 indicates that the innovative /w/ variants appear more often than conservative variants, in both positions. However, the initial-position shows a higher rate of innovative use than the medial-position. As shown in Figure 4.2, 70% of /w/ initial-position variants are innovative. This drops to roughly 50% in medial position. This result is consistent

with the prediction that /w/ in initial position is more likely to be adapted using innovative variants as compared to medial /w/ in medial position.

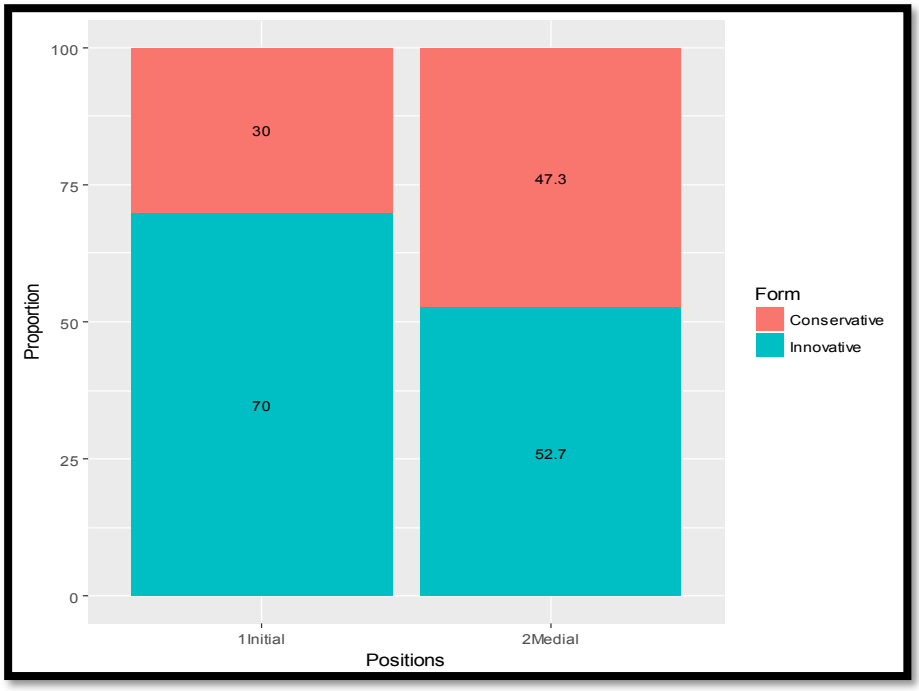

Figure 4.2: Positional distribution for /w/ variants

4.2.4 Gender

The distribution of innovative and conservative forms for male and female authors are shown in Figure 4.3. As we see, the innovative forms occur almost identically among male (66.5%) and female (65.8%) writers. This result is inconsistent with our predictions where we expected female authors would favour the conservative variants more than male writers (Tagliamonte, 2011) when borrowing English words containing /w/ sequence(s).

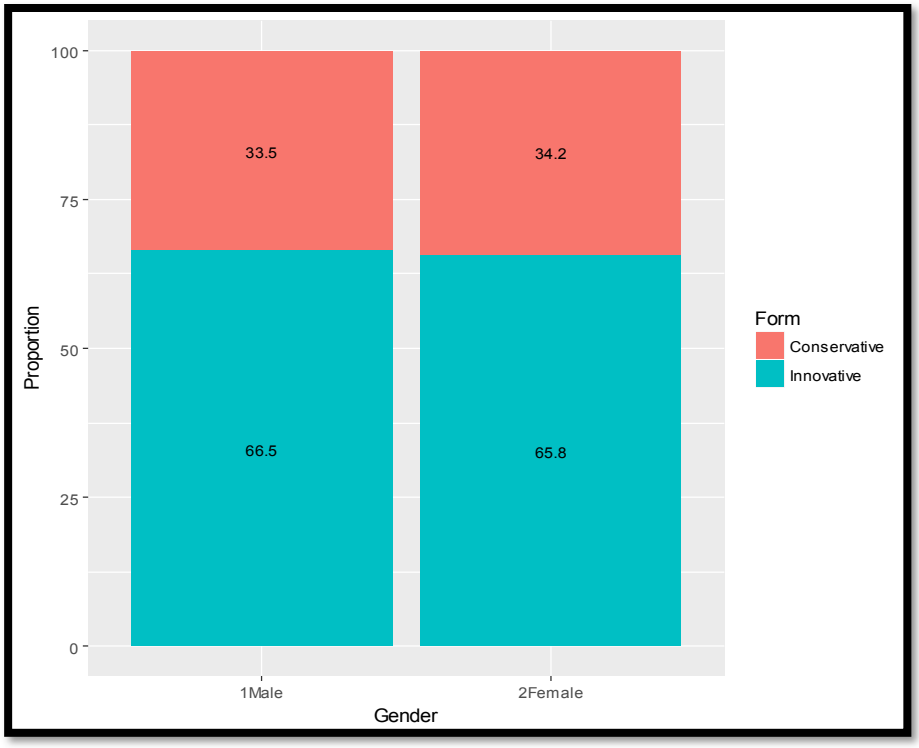

Figure 4.3: The distribution of /w/ variants for Males and Females

4.2.5 Age (YOB)

Descriptive statistics for age are shown in Figure 4.4. Here we see an overall increase in the usage of innovative variants from the 1910's to the 1980's. Surprisingly, the youngest writers in the dataset, those born in the 1990's, seem to use the novel variant less often (66.3%) than their counterparts from a decade earlier. Another unexpected finding is that those born in the first decade of the 20th century (1900's) were found to use innovative forms at a rate of 45.5%, much higher than those born in the 1910's (27.8%) and 1920's (44.6%).

The cohort of writers born in the 1800's, indeed, is the combined result for people who were born in the 1860's, 1870's, 1880's, and 1890's. This group of writers is the oldest group, and as predicted, older people use mainly conservative variants (Rebuck, 2002; Dougill, 2008; Barrs, 2011; Irwin, 2011; BKK, 2012; Inagawa, 2015; Kubozono, 2015a & 2015b), as much as 83.3%, in their writings.

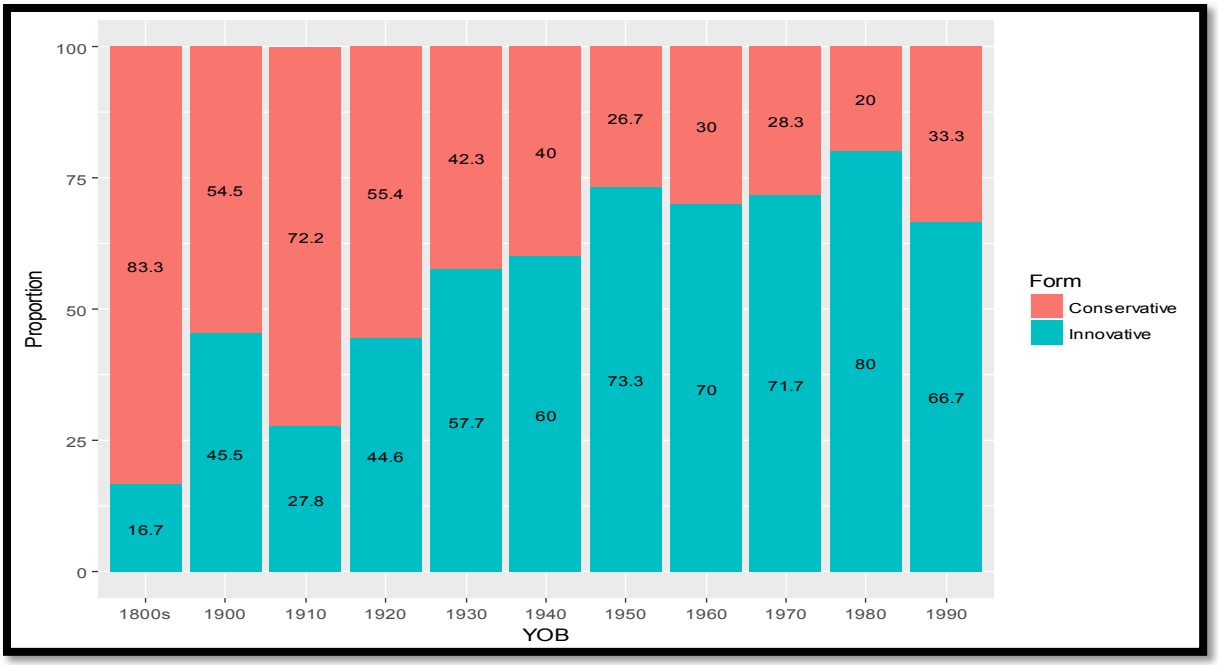

Figure 4.4: The distribution of /w/ variable for Age (YOB)

4.2.6 Registers

We have seen the results for both linguistic factors and author demographics. Let's turn now to look at the effect of the written register on stylistic variation. Figure 4.5 shows that innovative forms occur more often than conservative forms in all genres for the /w/ variables. For example, the innovative forms occur in the Book register at a rate of 67.7%, followed by 63.3% for Publication, and 63.0% for the Other category which includes governmental and legal genres. However, it was predicted that the Web register would have the most use of innovative katakana, which is indeed the case (72.1%).

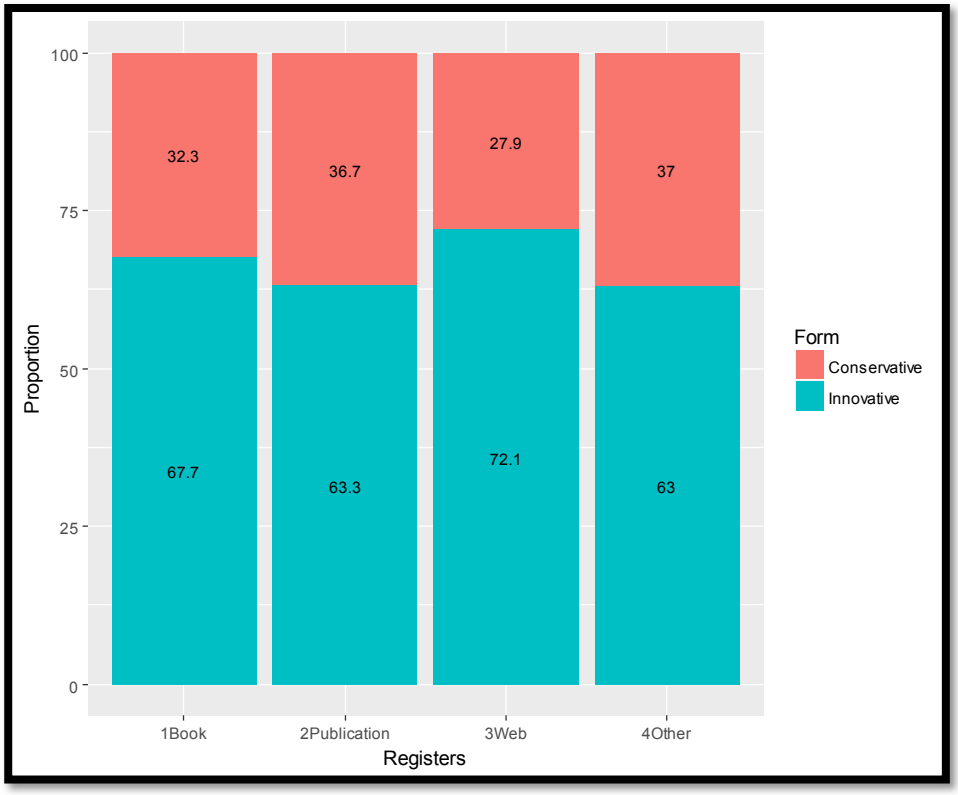

Figure 4.5: The distribution of /w/ variable across Registers

4.2.7 Publication Year

Figure 4.6 shows the use of conservative and innovative variants for different decades ranging from the 1970's to the 2000's. The distributional pattern shows a positive relationship between publication year and innovative orthography. Across the four decades, the 1970's are the only years where innovative variants make up the minority. Japanese writers tended to favour conservative forms (64.0%). The other three decades show innovative variants to occur at higher rates.

.

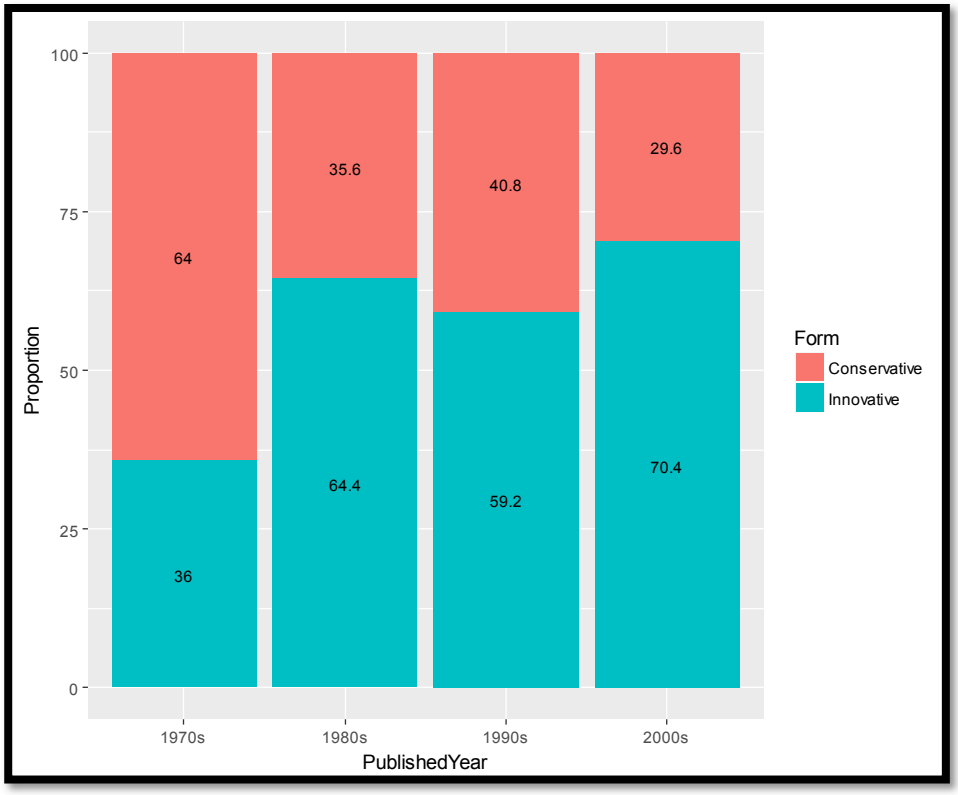

Figure 4.6: The distribution of /w/ Variable for Published Year in decades

4.2.8 Multivariate Analysis

This study used logistic regression[26], one of the generalized linear models, to examine the correlation between the dependent factor (the use of the innovative form) and the independent (both internal and external) factors such as position, gender, age, types of register, and publication year. The results of the multivariate analysis are presented in Table 4.3. I will present each result in turn.

The probability of any of the factors being is found in the p-value or Pr(>|z|) column. For example, the effect of Gender on the realization of innovative forms is statistically non-

[26] The function applied in R: model = glm(formula = FORM ~ POSITION + GENDER + YOB + REGISTER + PUBLISHED.YEAR, family = binomial(link = "logit"), data = w)

significant (as evidenced by a p-value of 0.159). All other factors were found to be significant. However, because some of the estimates are very small numbers, meaning that effect sizes are small, I have converted the estimates to log odds or Odds ratio[27], which are easier to interpret. These are found in the final column of Table 4.3.

Table 4.3: The statistical results of multivariate analysis for /w/ variables

	Estimate	Std. Error	z value	Pr(>|z|)	Odds Ratio
(Intercept)	-121	9.83	-12.32	<0.0001	<0.0001
Linguistic Factor					
Position-Medial	-0.766	0.0646	-11.85	<0.0001	0.465
Social Factors					
Gender-M	0.116	0.0824	1.41	0.159	1.12
Age (YOB)	-0.0002	<0.0001	-5.33	<0.0001	1.00
Register-Other	0.663	0.132	5.02	<0.0001	1.94
Register-Publication	-0.543	0.0725	-7.49	<0.0001	0.581
Register-Web	-0.187	0.0746	-2.50	0.012	0.830
Publication Year	0.0610	0.0049	12.43	<0.0001	1.06

The factor position, with the baseline of Initial position, was found to have a statistically significant effect on the choice of innovative variants. In particular, when the variable is found in Medial position, the odds of finding an innovative variant is 47% as likely as in Initial position. As predicted for the /w/ variable, this statistic shows that Initial Position, relative to Medial position, favours innovative forms.

Surprisingly, the age factor has an odds ratio of 1.0 meaning that there is no variation in the use of innovative variants between old and young writers. This result is not consistent with the descriptive result in Figure 4.4. This inconsistency is due to the fact that all data for /w/ variable were considered in this multivariate regression, but the descriptive analysis as presented in Figure

[27] As 1.0 is the value associated with no effect at all relative to the baseline, anything above 1.0 is a positive effect and any value below is a negative effect.

4.4, considered a relatively smaller dataset (less than 30% as recorded in Table 3.5) of the overall data for the /w/ variables.

Publication type was also found to affect the use of innovative katakana. Book functioned as the default level of register in the model. Interestingly, texts written for the Web are less likely to contain novel instances of /w/ (83% relative to Books). Likewise, the register labeled Publication is 58% as likely to exhibit innovative forms compared to Book. Innovative variants are almost twice as likely (odds ratio = 1.94) in the register of Other relative to Book register. This positive relationship was unexpected as the descriptive results presented above in Figure 4.5 showed no evidence of it. This may be because the distribution of innovative and conservative variants across all registers in Figure 4.5 is examined without considering the Publication Year factor. So, if considering the information of Publication Year for all registers, the descriptive results are different, specifically for the Other genre as shown in Figure 4.8. Compared to Figure 4.7, the baseline (Book) distribution, the register of Other shows that the usage of innovative variants over time is approximately double for most of the years. This comparative observation is compatible with the positive effect discussed above.

Finally, Table 4.3 shows that Publication Year has a positive effect with an odds ratio of 1.06 showing a 6% increase in the likelihood of innovative variants per year.

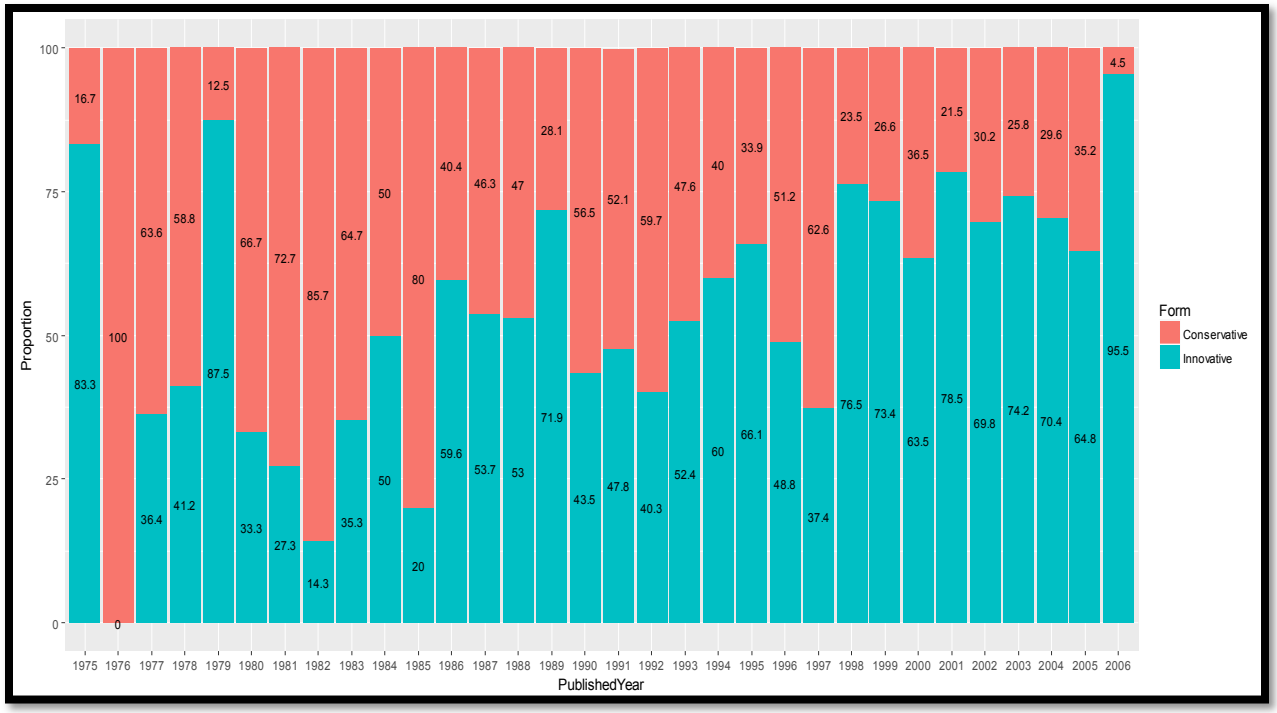

Figure 4.7: The distributional result of Book (baseline) register for /w/ variable from 1975 to 2008

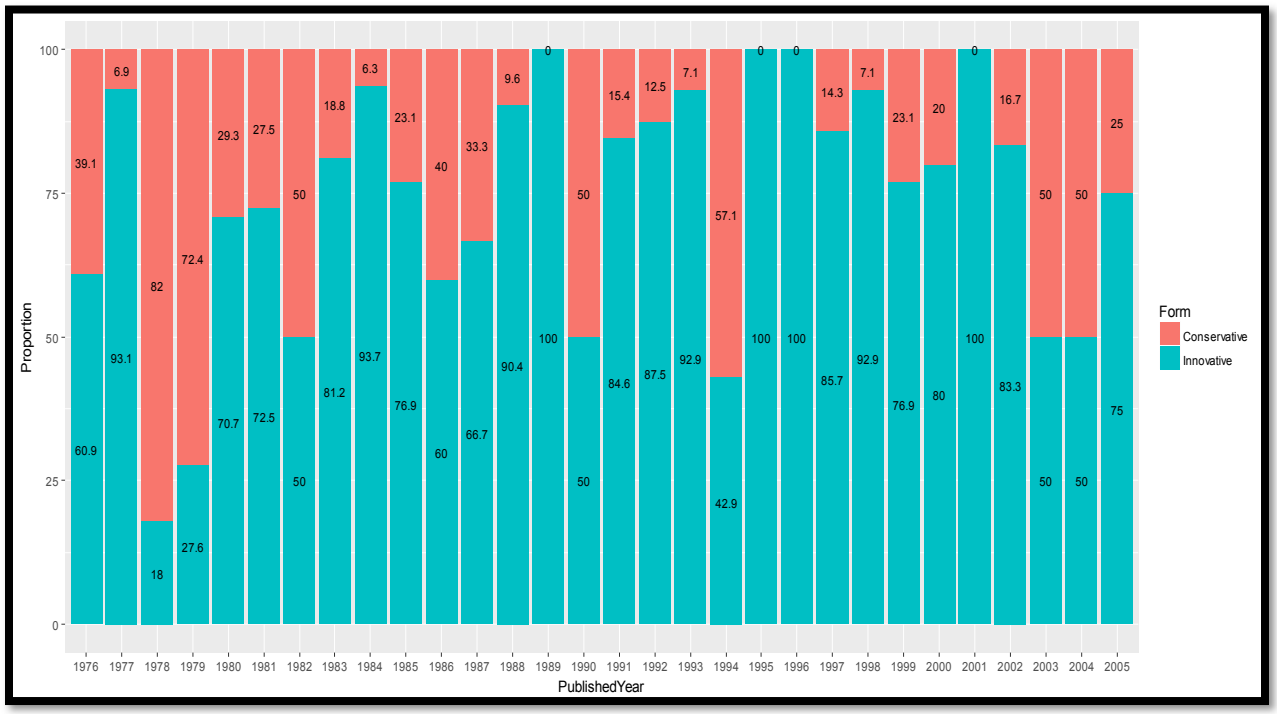

Figure 4.8: The distributional result of Other register for /w/ variable from 1975 to 2005

4.3 Analytical Results for /v/ Variable

In general, the descriptive and statistical findings regarding the /v/ variable differ substantially from those patterns found in the data for the /w/ variable. Most notably overall usage of innovative forms for /v/ is much lower than what is found for /w/. I report the /v/ results according to the same organization as /w/ in section 4.2.

4.3.1 Orthographic of /v/ Variants in Japanese

Similar to the orthographic finding observed for the /w/ variable, there is a loanword containing /v/ which also shows unanticipated orthography in Japanese. This sole exception among all /v/-based loanwords is the word “vodka”, in (4.1), which is conventionally realized as conservative /uokka/ and innovative /wokka/ instead of */bokka/ and */vokka/. At least in this corpus, I found no token for those anticipated variants. There are some rationales behind this finding that I further discuss in the next chapter.

	Donor word	**Conservative**	**Innovative**
(4.1)	vodka	＊ボッカ*/bokka/ ウオッカ/uokka/	＊ヴォッカ*/vokka/ ウォッカ/wokka/

4.3.2 The Absence of Native /v/ Sequences

As discussed in the literature review, the Japanese phonological system lacks the /v/ phone. So, as anticipated, the search results for /v/ variables consist of a very significant percentage of 95.9% of conservative variants and only 4.1% for innovative variants as shown in Figure 4.9.

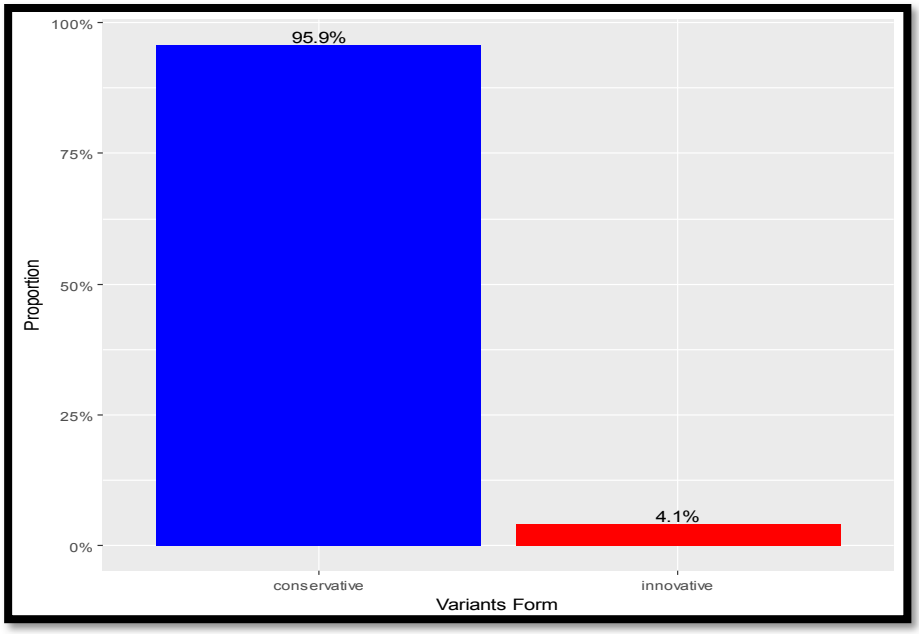

Figure 4.9: The overall distribution of /v/ variables

4.3.3 Positions of /v/ Variables

Unlike /w/ sequences, /v/ sequences can appear in final position, in addition to word-initial and word-medial position. The distributional result for /v/ sequences is shown in Figure 4.10.

The conservative proportions in Figure 4.10 show that medial-position variants are most likely to be conservative, 98.12%, followed by the initial-position variants (92.88%) and the final-position variants with the smallest probability of 91.02%. In other words, innovative /v/ variants occur most likely at the final and initial positions (8.98% and 7.12% respectively) as compared to the medial-position where innovative variants have the least probability value of 1.88%. This descriptive result is compatible with the hypothesis made in Lively et al.'s (1994) finding which says final /v/ variables should be more likely to adapt as innovative variants than initial /v/ variables, and medial /v/ variables should show the least tendency being innovated.

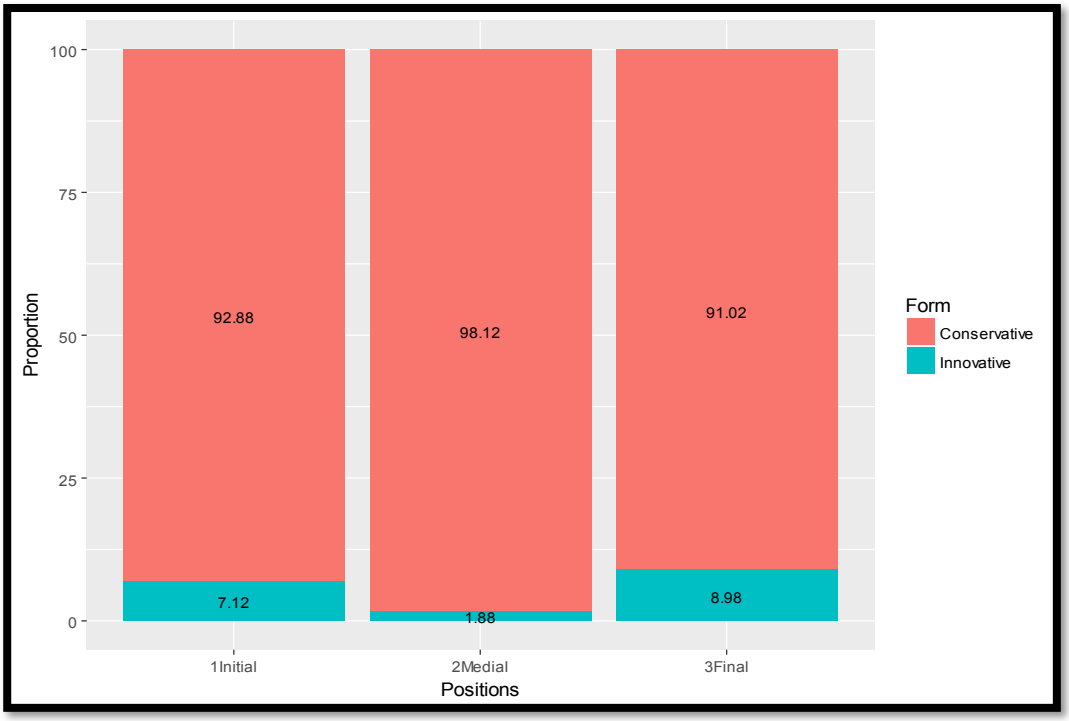

Figure 4.10: Positional distribution for /v/ variants

4.3.4 Gender

The likelihoods of men and women using conservative /v/ variants are shown in Figure 4.11. Figure 4.11 shows that female and male writers have a high tendency to opt for conservative /v/ variants when adapting the /v/ variables from English. Both genders show more than 90%, with 92.91% in males and 94.90% in females.

In contrast to the gender factor for /w/ variables, the finding for /v/ variables is considered compatible with the hypothesis that males are more likely to utilize innovative variants than females (Trudgill, 2000; Tagliamonte, 2011).

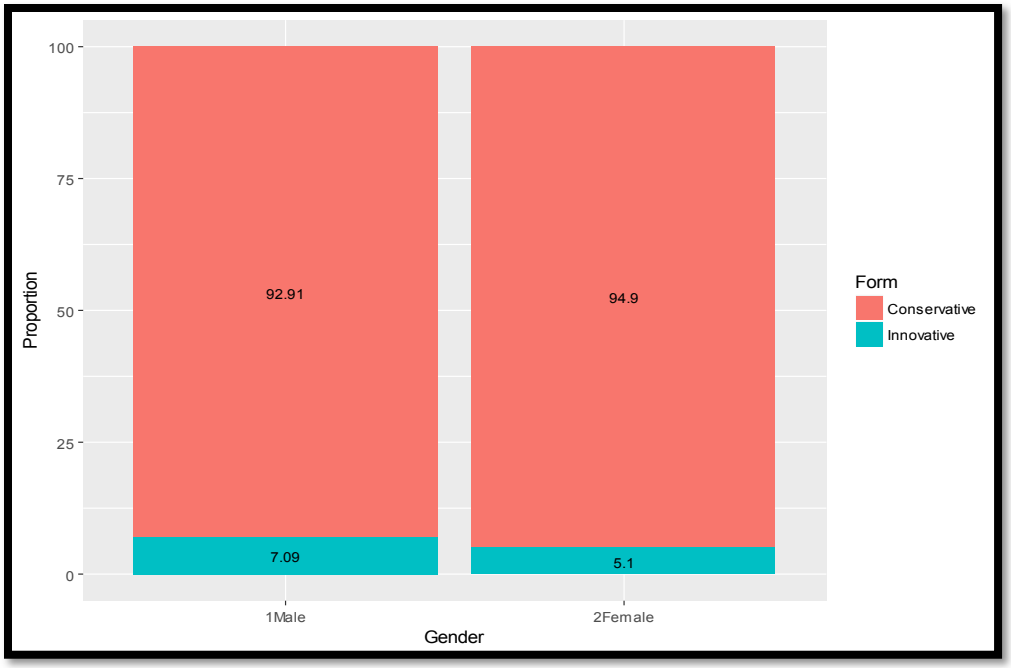

Figure 4.11: The distribution of /v/ variables for Males and Females

4.3.5 Age (YOB)

The usage of both innovative and conservative variants for different age cohorts is presented in Figure 4.12. To my surprise, two of the youngest groups of writers, those born in the 1980's and the 1990's have 0% use of innovative /v/ variants. Another surprising finding is that some of the older groups, i.e., those born in the 1900's and the 1910's, show higher usage of innovative forms in their writings than the rest of the younger groups.

This indicates that younger writers opt solely for conservative over innovative variants. This finding is odd and conflicts with the prediction of this study, i.e., younger people should show more accommodation of novel adaptation than older cohorts (Rebuck, 2002; Dougill, 2008; Barrs, 2011; Irwin, 2011; BKK, 2012; Inagawa, 2015; Kubozono, 2015a & 2015b).

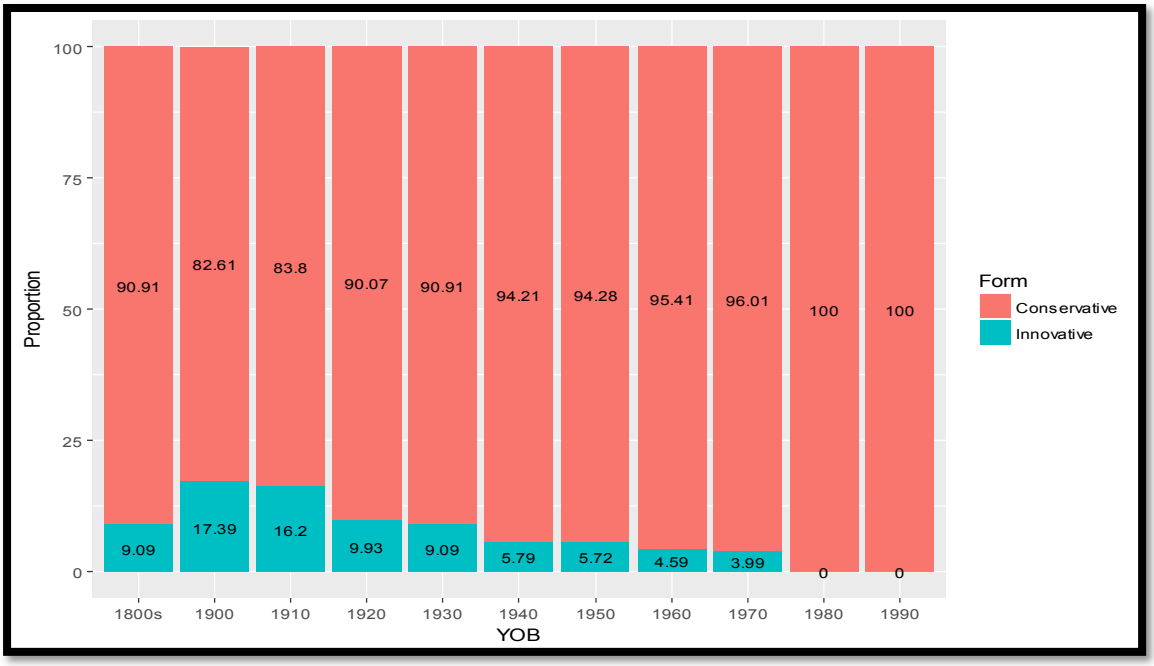

Figure 4.12: The distribution of /v/ variable for Age (YOB*)*

4.3.6 Registers

All registers in Figure 4.13 show relatively high probabilities, more than 90%, of using conservative variants of /v/. The register of Book shows the highest percentage (6.44%) of innovative adaptation, followed by Web (2.81%), Publication (1.64%), and lastly the Other register which has the minimum novel usage (0.33%) across the registers.

The result in Figure 4.13 indicates that Book has unexpectedly contributed more innovative variants as opposed to Web which was hypothesized to utilize more novel variants in this internet-based register (Sakamoto, 2002).

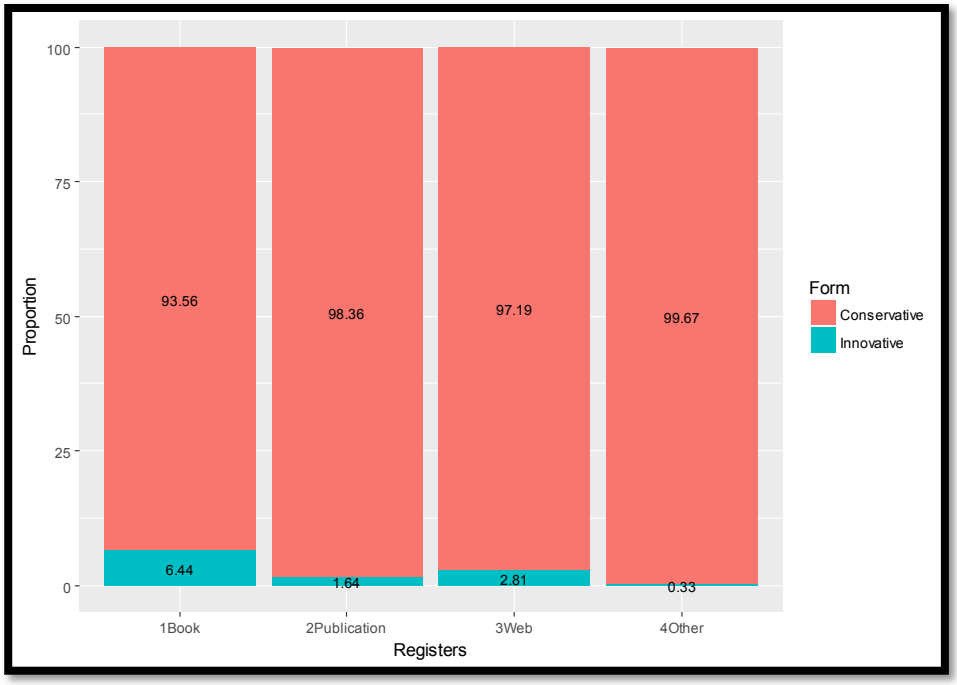

Figure 4.13: The distribution of /v/ variable across Registers

4.3.7 Publication Year

Figure 4.14 records the innovative usage in publications in the corpus from the 1970s to the 1990's. Even though the percentages of innovative variants remain less than 10%, the use of innovative forms has increased over time. The exception to this pattern is the drop in the most recent decade, the 2000's. The conservative variants are favoured again after the 1990's as shown in the data from Figure 4.14: the likelihood is increased to 96.37% (in the 2000's) from 93.22% (in the 1990's). This also means that the innovative variants become less popular with the Japanese writers for recent publications.

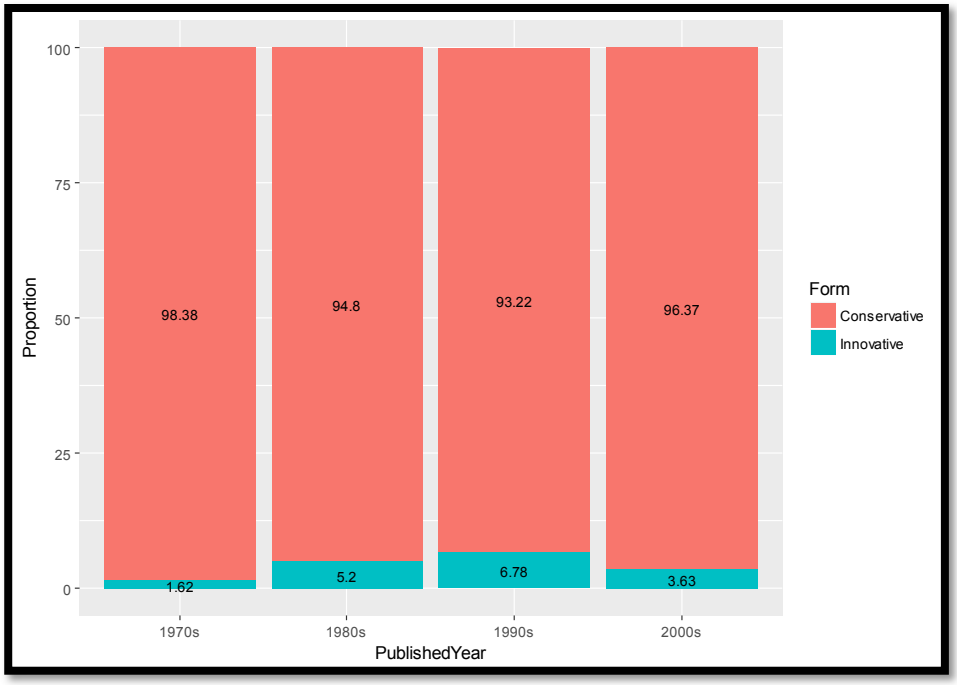

Figure 4.14: The distribution of /v/ variable for Published Year in decades

4.3.8 Multivariate Analysis

I ran the same generalized linear regression[28] for the subset as of the data containing the /v/ variable I did for the /w/ variable, to test the influences of social and linguistic factors on the use of innovative /v/ variants in the written texts in the corpus. I presented the results from multivariate analysis of /v/ variables in Table 4.4.

According to the p-value or Pr(>|z|) column, the effects for most of the factors are significant with the exception of Final position. Relative to the baseline of Initial position, the effect of Final position is insignificant as its p-value is 0.387.

[28] The function used in R: model = glm(formula = FORM ~ POSITION + GENDER + YOB + REGISTER + PUBLISHED.YEAR, family = binomial(link = "logit"), data = v)

Table 4.4: The statistical results of multivariate analysis for /v/ variables

	Estimate	Std. Error	z value	Pr(>|z|)	Odds Ratio
(Intercept)	0.959	8.48	11.31	<0.0001	$4.61e^{+41}$
Linguistic Factor					
Position-Medial	-1.38	0.0786	-17.57	<0.0001	0.251
Position-Final	0.0635	0.0733	0.87	0.387	1.07
Social Factors					
Gender-M	0.223	0.0720	3.09	0.00199	1.25
Age (YOB)	-0.0002	<0.0001	-5.31	<0.0001	1.00
Register-Other	-2.92	0.235	-12.42	<0.0001	0.0542
Register-Publication	-1.03	0.0864	-11.90	<0.0001	0.357
Register-Web	-0.279	0.0749	-3.73	0.0002	0.756
Publication Year	-0.0490	0.00424	-11.54	<0.0001	0.952

Similar to the /w/ variable, I discuss the nature of the effects below with reference to the odds ratios. For example, the effect of Medial position is significant and strong, with a log odds ratio of 0.251. This means that the likelihood of an innovative variant occurring in Medial position is approximately 25% as likely as in the baseline of Initial position. This difference between medial and initial positions for the /w/ variable is consistent with the hypothesis that innovative usage will be more frequent in initial position. However, the insignificant result for Final position shows that there is no significant difference between usage of innovative forms in final position relative to the baseline of initial position.

With respect to gender (Female as the baseline), Male has a log odds ratio of 1.25 meaning that a male writer is 1.25 times as likely to produce an innovative variant as a female writer. This statistical data is consistent with the descriptive result in Figure 4.11 where males have a slightly higher rate of innovative /v/ usage, i.e., 7.1% male as opposed to 5.1% female.

As was the case with the /w/ variable, there is also no effect of age on the use of innovative forms in the representation of the /v/ variable. The odds ratio is 1.0 which indicated that the

likelihood of using an innovative variant is unchanged across different age cohorts. This result is incompatible with the expectation for age factor, as discussed previously in relation to the /w/ variable.

Concerning the types of genres with the baseline of Book register, all registers show significant negative effects on use of innovative forms. In particular, the register of Other shows the strongest effect. The register of Other has a log odds ratio of 0.054 meaning that the occurrence of an innovative variant is only 5.4% as likely in the data from the Other register compared to the baseline (Book). This significant negative effect of the Other register on the use of innovative katakana is consistent with predictions as the genres in this register consist of formal publications, namely governmental and legal documents.

There are also negative effects for Publication and Web. For Publication, the value of 0.36 indicates that use of innovative variants is 36% as likely in Publication relative to Book. Register Web has a log odds ratio of 0.76 meaning that the occurrence of an innovative variant in web data is 76% as likely as in the baseline case of Book. So, Web is still a negative effect, but not as strong as Publication and Other. The negative effects for all register categories shown in Table 4.4 indicate that Book register favours the use of innovative forms. This is consistent with the descriptive finding in Figure 4.13.

The factor of Publication Year also has a negative effect with an odds ratio of 0.95. This means that there is a 5% decrease in the likelihood of an innovative variant being used for every year.

4.4 Summary

The descriptive and multivariate findings presented in this chapter indicate that English-origin loanwords containing the /w/ variable are more likely to be written in Japanese with innovative variants as opposed to words containing the /v/ variable. This finding is consistent with the first linguistic hypothesis which is that the presence of native /w/ phones in the native Japanese phonology will favour the use of innovative forms for /w/ variables. The results from this study have demonstrated that the existence of native sequences of /wa/, /wu/, and /wo/ motivates Japanese writers to use novel katakana variants of /we/, /wi/, and /wo/. In contrast with the /w/ results, the lack of native /v/ sequences in the Japanese orthographical and phonological systems may contribute to the overall low rate of innovative forms for the /v/ variable.

With respect to position, medial position has a significant negative effect relative to initial position for both /w/ and /v/. This indicates that both variables are very likely to be innovated in initial position as compared to medial position. For /v/, which also occurs in final position, both initial and final positions were more likely to show use of innovative variants than medial position but there was no significance difference between innovative usage in initial and final positions. These findings are considered consistent with the predictions that medial position should disfavour the use of innovative forms. The predicted difference between initial and final positions was not found for the /v/ variable.

Looking at social factors, we see that the preference for using innovative /w/ and /v/ variants is not determined by age. Gender has an effect on the use of /v/ innovative variants where male authors are 25% more likely to use novel /v/ variants than female authors. For the /w/ variable, on the other hand, there is no significant effect of gender.

Considering the genre of publications, there are some unexpected findings in the use of innovative variants. For example, the register of Web, which was expected to show the highest number of innovative variants, showed a negative effect on use of innovative /w/ and /v/ variants relative to the register of Book. The most formal type of register, Other register, also presents unexpected results in the use of innovative /w/ and /v/ variants. Especially for innovative /w/ usage, results show that the Other register, which includes governmental and legal genres, has the greatest usage of innovative forms across the registers.

Lastly, both descriptive and multivariate results show that innovative orthography for the /w/ variable has a positive effect whereas the /v/ variable has a negative effect over time. This means that the /v/ variable once again shows an unexpected result which is inconsistent with the hypothesis of expecting use of innovative variants to be increasing diachronically.

After reporting the results in this chapter, I further discuss the interpretation of these findings in the following chapter.

CHAPTER 5: DISCUSSION

This chapter discusses possible interpretations of the results in the previous chapter with respect to the influence of linguistic and sociolinguistic factors on use of innovative katakana. I also compare the findings of this study with other research and discuss unanticipated outcomes of this study.

The discussion starts with irregularities, such as the exceptional orthographic variation found in “sandwich”, “whisky”, and “vodka” (section 5.1.1). Then, I present general and more principled findings in sections 5.1.2 and 5.1.3. In section 5.1.2, I review the effect of the presence and absence of native sequences on the use of innovative variants practice in the writing of Japanese. After that, I discuss, in section 5.1.3, the positional influence on the use of innovative variants and consider some reasons for the high rates of innovative variants in initial position relative to medial position.

Section 5.2 discusses findings regarding social factors, namely gender, age, register, and publication year.

5.1 Linguistic Factors

5.1.1 Exceptional Orthographic Forms

The results from sections 4.2.1 and 4.3.1 show that a few loanwords have an extra variant other than their expected conservative and innovative variants. For the /w/ variables, the loanwords of “sandwich” and “whisky” show exceptional results as recorded in Table 4.2, because they have more variants than the conservative /ui/ and innovative /wi/ variants found for other lexical items.

Apart from the conservative form of /sando**ui**ʧi/ and the innovative form of /sando**wi**ʧi/, "sandwich" has an extra variant of /sando**i**ʧi/. The presence of this variant in the corpus may follow from the early time of borrowing of this particular word. The sequence of /wi/ did not start appearing in dictionaries until the 1950's (Stanlaw, 2002). So, without the option of this innovative katakana character, /i/ is the closest sequence to donor /wɪ/ for this word. Although /i/ is an old replacement for foreign /wɪ/[29], the /i/ variant of "sandwich" is highly integrated and embraced by the Japanese. This is because this study found that /i/ is still extensively used in the most recent publication in the corpus, from 2000 to 2008 (Appendix 6). This preservation of /i/ in "sandwich" suggests that the Japanese obey the recommendations from the corresponding prescriptive authorities (Monbusho, 1991; Bunkacho, 1991; NINJAL, 2003a & 2003b & 2004 & 2006; JTCA, 2015) as one of the recommendations from them is to retain the historical katakana forms that are intelligible to most of the Japanese (Irwin, 2011).

Nonetheless, one of the variants in "sandwich" which is found in 1975 (the earliest year in the corpus), as shown in Appendix 5, is the innovative variant of /wi/ (87 cases or 13.55%)[30], instead of /ui/ (5 cases or 0.78%) or /i/ (550 cases or 85.67%). Surprisingly, the conservative variant of /ui/ for "sandwich" only appears 5 times in the corpus in very recent publications, in 2005 and 2008 (Appendix 7). Furthermore, this so-called conservative /ui/ variant is only found in the register of Web, which is considered as modern and informal. Perhaps, "sandwich" is a loanword that does not follow the anticipated pattern of other loanwords as discussed in the following; instead "sandwich" has an elastic character which allows the /wɪ/ sequence in the donor

[29] The historical variant of /i/ can also be seen in the loanword of /suiʧi/ which is adapted from "switch" /swɪʧ/ (NHK, 2012).
[30] The total of cases found in this corpus for "sandwich" is 642.

form to be adopted flexibly as any of the three katakana variants of /i/, /ui/, and /wi/ in contemporary written Japanese.

For "whisky", the conservative variant of /ui/ is used from the 1970's to the 2000's and the innovative form of /wi/ is adapted from the 1980's to the 2000's. The data from the corpus show that both of these variants are used interchangeably for recent publications (Appendix 9). So, this example also suggests that some historical forms such as /i/ for "sandwich" and /ui/ for "whisky" persist despite the emergence of innovative katakana forms in Japanese. Like "sandwich", "whisky" also has an exceptional variant, namely /uwi/. The /uwi/ variant in "whisky" seems like an intermediate variant for the evolution from the conservative /ui/ variant to innovative /wi/ variant. This is because the /uwi/ variant is found in the governmental genre in 1978 and in library books in 1991, 1997, 1999, and 2003 (Appendix 8).

Exceptional variants can also be found in some lexical items containing the /v/ variables. For the loanword of "vodka", the English segment of /va/ is rendered as /uo/ and /wo/ as conservative and innovative variants as shown in (4.1), rather than with the expected /bo/ and /vo/ variants found in other words.

The interpretation of this interesting orthographic form for "vodka" relates to the perception of similar labial place features of non-native /w/ and /v/ sounds. The phonetics of /w/ and /v/ sound similar to Japanese listeners who do not have this contrast in their phonemic inventory. This is perhaps especially true when "vodka" was borrowed at a very early time, some time in between 1910 to 1927, as recorded in Irwin (2011, p. 46). Irwin (2011, p. 85) confirms that the voiced labiodental fricative /v/ in Russian words, such as "vodka", is usually adapted as /w/ or /u/ in Japanese. And this became the main reason to explain why the Japanese adapted "vodka" as /uokka/ and /wokka/ at the time of borrowing. Moreover, articulatory and perceptual factors may

play a heightened role at the time when these items were borrowed due to the relative lack of contact and exposure to foreign languages (Fais et al., 2005). Of the two exceptional variants found in "vodka", /wo/ and /uo/, /wo/ is a relatively novel variant, derived from bimoraic /uo/, which first appeared in the 1960's (Stanlaw, 2002). The result from this study also shows that "vodka" is an old loanword, with forms occurring in the corpus from 1978: /wotsuka/[31].

5.1.2 The Effect of Native sequences on Loanword Adaptation

The probabilities in Figure 4.1 reveal that the presence of native variants (/wa/, /wu/, and /wo/) in the Japanese system correspond to higher usage of innovative variants, 67.8%, as compared to 32.2% of conservative variants. The result in Figure 4.9 is also compatible with the expectations of this study as well because the proportion of conservative variants (95.9%) is vastly greater than innovative variants (4.1%) for /v/ variables in loanword adaptation. The rationale to explain the extensively low probability of novel variants in the corpus is the absence of /v/ sequence in the native inventory.

These findings from Figures 4.1 and 4.9 confirm that the presence of native consonants in the traditional inventory would favour the incorporation of the native consonant with other possible vowels to generate novel variants. These innovative variants of /wi/, /we/, and /wo/ are embraced by the Japanese with ease as compared to innovative /v/ sequences. There are some other explanations for the Japanese insisting on replacing English /v/ variables with conservative /b/ variants even if these /b/ variants would confuse audiences or readers. Firstly, this attitude is arguably influenced by the perception of the lack of /v/ in the traditional inventory system (Fais et al., 2005). Second is the katakana endorsement given by the Japanese bureaucracies (Monbusho,

[31] The donor word of "vodka" has six katakana variants as found in the corpus. The conservative variants are /uotsuka/, /uotoka/, and /uokka/. Their novel counterparts are /wotsuka/, /wotoka/, and /wokka/.

1991; Bunkacho, 1991; NINJAL, 2003a & 2003b & 2004 & 2006; JTCA, 2015) since 1955 to conserve bimoraic or conservative forms rather than monomoraic or novel forms (Irwin, 2011).

5.1.3 Positional Effect on the Use of Innovative Variants

The influence of the position within the word on the use of innovative variants is a relatively new domain in the research of innovative katakana in contemporary Japanese. Few, if any, studies have been done on the role of position in the use of innovative katakana forms. Some work on the effect of different positions is embedded in previous literature such as Lively et al. (1994). Their study examined the ability of Japanese speakers to differentiate between English /l/ and /r/. I adopted their findings to make the hypotheses for the effect of position in /w/- and /v/-based loanwords.

The frequencies recorded in Table 5.1 show that /w/ variables occurred more frequently in initial position overall and that there were more than twice as many innovative variants as conservative variants. However, for the smaller number of /w/ variables which occurred in medial position the number of conservative and innovative variants were close to equal (566 conservative forms and 631 innovative forms). This observed difference between initial and final position is supported by the statistical results in Table 4.3 which show that the /w/ variables in medial position are quite unlikely to be innovated compared to initial /w/ segments.

Table 5.1: The frequencies of both variants for /w/ and /v/ variables

	Initial	**Medial**	**Final**
/w/ conservative	**2519**	566	-
/w/ innovative	**5866**	631	-
/v/ conservative	21004	**35317**	2492
/v/ innovative	**1609**	678	246

Although the overall distribution of /v/ variables is substantially different from that of the /w/ variables, with a much smaller percentage of innovative variants occurring throughout the data set, the relative pattern of innovative forms with respect to initial and medial position is the same

as that which occurs with /w/, i.e., the initial position has a higher frequency of innovative variants, with a smaller percentage of innovative forms occurring in medial position.

Unlike the /w/ variable, /v/ can also occur in final position. The descriptive finding in Figure 4.10 shows that the percentage of using innovative forms in final position is higher than that in initial and medial positions. This descriptive result is consistent with the hypothesis, but this difference is not significant in the multivariate analysis (Table 4.4) which shows that final-position has no effect in the innovative usage of /v/ variants.

These findings, in sum, reveal that initial /w/ and /v/ segments are very likely to appear as innovative variants relative to medial /w/ and /v/ segments. Although the multivariate statistic for /v/ in final position was inconsistent with the prediction, final position nonetheless favours use of innovative forms relative to medial position. Below, I present some /v/-based donor words to illustrate the tendency for segments in initial and final positions to be innovated as novel variants as opposed to medial segments.

	Donor word	**Position**	**Conservative**	**Innovative**
(5.1)	van /van/	Initial	/ban/	/van/
(5.2)	bun /bən/	Initial	/ban/	-
(5.3)	motivation /moʊtəveɪʃən/	Medial	/moʧibeʃon/	/moʧiveʃon /
(5.4)	curve /kərv/	Final	/ka:bu/	/ka:vu/
(5.5)	curb / kərb/	Final	/ka:bu/	-

The medial /v/ segment in “motivation” (5.3) is not that obviously distinguished from the donor segment when the conservsative /be/ is used instead of the novel /ve/. This is because Japanese readers could still interpret the donor word if they listen to or look at the preceding and following segments in “motivation”.

As compared to medial /v/ segments, initial and final /v/ segments have a higher possibility to be perceived incorrectly especially for donor words having the same katakana forms. For

example, the English word of “bun” in (5.2) has the same katakana form with the conservative form of “van”, i.e., /ban/ in Japanese. Also, readers might misinterpret the conservative form of “curve”, /ka:bu/ in (5.4) as another English word, such as “curb” shown in (5.5). The homogenous initial /b/ form for “van” vesus “bun”, and “curve” versus “curb” could make the Japanese readers misinterpret “van”/van/ as “bun”/bən/ and vice versa. These homogenous forms also reflect the importance of using the innovative katakana variant particularly in initial position so that Japanese speakers and readers can distinguish the appropriate donor words from foreign languages. The examples of conservative forms for “van” and “bun” in (5.1) and (5.2) also show explicitly why the initial segments are more likely to be innovated than the medial segments because the initial segments are more perceptually salient than the medial ones.

The effect of using the innovative form in different positions as discussed here indeed support Lively et al.’s (1994) findings that Japanese could recognize better the final and initial English /l/ and /r/ as opposed to medial /l/ and /r/. Therefore, the innovative katakana forms should be used more frequently especially in initial and final positions in order to assist Japanese readers to distinguish the appropriate donor words from foreign languages.

5.2 Sociolinguistic Factors

5.2.1 The Role of Gender

Although I expected males to favour the use of innovative variants relative to females, for the /w/ variables (Figures 4.3), I found an unanticipated result: Males (66.5%) and Females (65.8%) have the same preference in their writing behaviours, i.e., they prefer novel variants for /w/ segments. This finding is inconsistent with the literature by Trudgill (2000), Labov (2001), and Taglimonte

(2011). The female writers did not favour the conservative forms of /w/ variants instead they use innovative /w/ variants as often as the male writers. The possible interpretation of this writing behaviour is the presence of native sequences of /wa/, /wu/, and /wo/ which encourage Japanese women to embrace the innovative variants of /we/, /wi/, and /wo/ and use them in their writings.

With respect to the /v/ variables, the usage of innovative /v/ variants is slightly different between men and women writers. The probability of male writers (7.09%) to use innovative /v/ variants is higher than female writers (5.10%), as shown in Figure 4.11. As compared to the /w/ result, the /v/ result is relatively compatible with the hypothesis that women tend to use more prestige or conservative forms (Tagliamonte, 2011) than men.

5.2.2 Age

If considering the descriptive results from Figure 4.4, the usage of innovative /w/ variants shows an increase in use from those born in the 1910's to the 1980's. However, the youngest group of writers, those born in the 1990's, show less preference in using innovative /w/ variants as compared to older groups. This distribution of /w/ variables according to age is not consistent with the expectation that younger people are the generations who are importing and using innovative forms as indicated by Rebuck (2002), Dougill (2008), Barrs (2011), Irwin (2011), BKK (2012), Inagawa (2015), and Kubozono (2015a & 2015b).

For /v/ variables, the distributional result in Figure 4.12 also shows that the usage of innovative /v/ variants is disfavoured by younger writers as opposed to older writers. This finding is also incompatible with the hypothesis that younger writers use innovative variants more frequently. This might be a case of age-graded variation (Labov, 1994; Matsuda, 2003; Tagliamonte, 2011) as the influence from katakana guidelines recommended by Monbusho (1991),

Bunkacho (1991), NINJAL (2003a & 2003b & 2004 & 2006) and JTCA (2015) after 1991 has urged these millennials to conform to the authorities' main recommendation: use traditional Japanese katakana forms whenever possible. To be specific, this writing behaviour among the writers who were born in the 1980's and the 1990's is because of the need to use prestige forms in work environments (Matsuda, 2003; Tagliamonte, 2011). These young writers are conscious of those katakana recommendations, and they preferred to use more conservative forms than innovative forms.

The statistical analyses for /w/ and /v/ indicates that there is no effect of age on the use of innovative katakana. The odds ratio values from the logistic regression analyses are 1.00 (Tables 4.3 and 4.4) showing that increase in age has no impact on katakana usage. Perhaps these unpredicted findings for /v/ and /w/ are affected by the large size of the "0" group where the age of the author was not known, or multiple authors of different ages that also contributed to the analyses. In sociolinguistic studies, the social factor of age is often interpreted as providing evidence regarding the change in use of a variable over time. Usage patterns shown by older speakers are interpreted as representing the pattern of usage that was common among young speakers at the time that those speakers were young themselves, according to the apparent-time hypothesis (e.g. Labov, 1972). Fortunately, the corpus data considered in this study provides another source of information regarding change over time in the form of publication year.

5.2.3 Publication Year

Unlike the factor of age, data on Publication Year[32] is available for all data downloaded from the corpus. So, the descriptive results of Publication Year in Figures 4.6 and 4.14 are suitable to use for interpreting the possible real-time change in the use of innovative variants.

Considering the descriptive result in Figure 4.6 for /w/ variables, publications in the 1970's use less innovative /w/ variants (36.0%) and the recent publications have more innovative variants (70.4%). Similarly, the result from Figure 4.14 also shows a slight increase from the publication in the 1970's (1.62 %) to the publications in the 2000's (3.63%). These findings are interpreted as showing an increase in use of innovative variants over time.

The diachronic change in use of innovative variants also can be interpreted from the descriptive results of Publication Year in Figures 4.6 and 4.14. As anticipated in section 2.5.6, the innovative /w/ usage has increased from 36.0% to 70.4% and this is foreseeable because NHK (2015) urges writers to use innovative /we/, /wi/, and /wo/ in recent loanwords, such as /webu/ for "web" even though the conservative /ue/, /ui/, and /uo/ are still acceptable for the fused loanwords.

However, the distributional result for /v/ variants shows that the use of innovative /v/ variants has increased from the 1970's until the 1990's and then decreased from 6.78% to 3.63% in the 2000's. The multivariate statistic of the use of innovative /v/ variants also shows a negative effect which indicates that the probability of using innovative /v/ variant has decreased 5% since 1975. This finding is incompatible with the hypothesis that saying the usage of innovative forms should be increased over time. The possible reason for explaining the dispreference for using

[32] The multivariate statistical data of Publication Year in Tables 4.3 and 4.4 will not be discussed because of some missing years (2006 and 2007) due to the inadequacy of compiled resources in the corpus. It should be understandable because compiling data for a corpus is always time-consuming and so the size of a corpus "should or can be are changing all the time" (Baker, 2010, p. 12). This is also the reason for the Figures 4.6 and 4.14 are illustrated in the unit of decades.

innovative /v/ variants in recent years is the influence from the katakana guidelines by the Japanese divisions (Monbusho, 1991; Bunkacho, 1991; NINJAL, 2003a & 2003b & 2004 & 2006; JTCA, 2015) encouraging them to use the conservative forms especially after 1991 (Irwin, 2011).

5.2.4 Variations across different Registers

As hypothesized in this study, the register of Web comprises the highest rate (72.1%) of innovative variants for the /w/ variable (Figure 4.5). This observation seems consistent with the previous literature (Sakamoto, 2002; Tagliamonte, 2011) on the usage of novel variants in these online genres, or informal register[33]. However, the multivariate result in Table 4.3 shows the opposite interpretation, the Web register is only 83% as likely to find innovative /w/ variants as in the Book register. In other words, innovative variants are less likely on the Web although this register is considered less formal than Book.

Another unexpected result of the multivariate analysis is that the register of Other is 94% more likely to contain innovative /w/ variants relative to Book. This result is particularly surprising because the register of Other consists of "formal" genres such as Whitepapers, Legal Documents, and National Assembly Proceedings which were expected to prefer conservative /w/ forms in the publications. This finding conflicts with the prediction that people prefer the prestige forms (conservative forms in this case) in formal styles (Tagliamonte, 2011). This also reflects that the Japanese governmental authorities do not obey the innovative /w/ recommendation from Monbusho (1991), Bunkacho (1991), NINJAL (2003a & 2003b & 2004 & 2006), and JTCA (2015) which urge the Japanese people to conserve and utilize the conservative and intelligible forms as far as they could (Irwin, 2011).

[33] The internet or online genres are categorized as "informal" genres because the internet source in the corpus is compiled from *Yahoo! Blog* and *Yahoo! Answer*.

Regarding the descriptive results for the /v/ variables (Figure 4.13), the finding is counter to expectations as well because the register of Book has higher proportions of innovative forms (6.44%) than the register of Web (2.81%). But, the lowest rate of innovative forms is found in the Other register (0.33% in Figure 4.13). This shows consistency with the hypothesis that the formal-styled register should have more prestige or conservative forms (Sakamoto, 2002; Tagliamonte, 2011).

However, due to the constraint of different timelines for these registers, the multivariate statistic is more appropriate to use for further interpretations. As recorded in Table 4.4, the statistical result shows the Web register is 76% as unlikely to use innovative /v/ variants as the register of Book. This is consistent with the descriptive finding discussed above.

The multivariate result for the Other register shows a consistent finding with the descriptive results above because it is only 5.4% as likely to find innovative /v/ variants in the Other register relative to the Book register. In other words, the likelihood for innovative /v/ variants to occur in Other register (odds ratio = 0.0542) is lower than Publication register (odds ratio = 0.357) and Web register (odds ratio = 0.756). These statistics show that Japanese authorities adhere to the recommendations from their associates for the restricted use of innovative /v/ variants (Monbusho, 1991; Bunkacho, 1991; NINJAL, 2003a & 2003b & 2004 & 2006; JTCA, 2015).

Comparing the findings of both variables, we can see that innovative adaptation for the /w/ variable is embraced by Japanese speakers as opposed to innovative /v/ adaptation. The increase in use of innovative /w/ variants is not gendered and is found at high rates in governmental and legal genres. This contrasts with use of novel /v/ variants. The use of innovative forms for /v/ is not favoured by the Japanese authors and is found more frequent in less formal registers such as books and online texts. These contrastive findings also indicate that innovative /w/ variants have

gradually become recognized forms in Japanese whereas the novel /v/ variants still remain as stigmatized variants.

CHAPTER 6: CONCLUSION

In this final chapter, I evaluate the importance of these findings (section 6.1) and raise possible questions for future directions in this area (section 6.3). I also discuss the limitations of this research in section 6.2.

6.1 Overview of this Study

Following the linguistic interpretations (section 5.1) from the Discussion chapter, the /w/- and /v/-based loanwords from English can contain several possible orthographical variants: 1. Conservative (bimoraic) variants, 2. Innovative (monomoraic) variants, 3. Historical variants, and 4. Intermediate variants. For example, "sandwich" has not only the conservative and innovative forms but also the conserved historical variant of /i/ for donor /wɪ/. The /u/ variant is another historical variant for donor (Russian origin) /v/ segments such as /v/ in "vodka". In addition, this study shows that an intermediate form has appeared during the integration process of novel monomoraic forms for the word of "whisky" in Japanese. The intermediate form is the /uwi/ sequence which is derived from /u/ in the conservative /ui/ and combined with the innovative /wi/.

The findings from section 5.1.2 have confirmed that the presence of native /w/ sequences foster the adaptation and usage of innovative /w/ variants. And the absence of native sequences, for example /v/ sequences, inhibits the use of innovative /v/ variants. The /w/ and /v/ results are also compatible with the chart of innovative katakana characters which was arranged by Stanlaw (2002) according to their appearance from the 1950's to the 2000's: the consonants which correspond to native sequences in the traditional Japanese inventory were introduced in Japanese dictionaries earlier than those which do not (see Table 2.5 and Appendix C).

This study has also successfully provided some significant findings on the effect of position in the use of innovative variants in Japanese. As discussed in the interpretations presented in section 5.1.3, the initial /w/ and /v/ segments are more salient than the medial segments in the innovation of katakana forms. As compared to previous studies which emphasize katakana functions, recognition and comprehension, the frequency of use of distinctive loanwords (Kay, 1995; Rebuck, 2002; Olah, 2007; Daulton, 2004 & 2008; Dougill, 2008; BKK, 2010 & 2012; Inagawa, 2015; Daulton, 2015) and other (non-) phonological loanword adaptations (Tsuchida, 1995; Stanlaw, 2002; Smith, 2006 & 2009; Crawford, 2008; Kochetov, 2008; Paradis & LaCharité, 2011; Irwin, 2011; Shoji & Shoji, 2014; Pinter, 2015; Kubozono, 2015b), this finding is relatively novel for linguistic studies of the (innovative) katakana forms in Japanese.

In light of the sociolinguistic analysis, most of the findings are not consistent with established principles of variation and change. First, for the factor gender, it was anticipated that women would use more conservative forms as opposed to men (Trudgill, 2000; Labov, 2011; Tagliamonte, 2011). But for the /w/ variables, there is no gender effect on the innovative usage because the female writers favour the innovative /w/ variants as much as the male writers do. For the /v/ variables, male authors show a slightly higher preference in using innovative /v/ variants than women authors.

Concerning the factor of age, the young(er) writers were expected to use more innovative variants than the older writers (Rebuck, 2002; Dougill, 2008; Barrs,2011; Irwin, 2011; BKK, 2012; Inagawa, 2015; Kubozono, 2015a & 2015b). However, the descriptive findings (Figures 4.4 and 4.12) show that the younger writers disfavour the innovative /w/ and /v/ variants as compared to the older cohorts. The multivariate statistical findings also indicate that there is no age effect on the usage of innovative /w/ and /v/ variants. Even though the descriptive results of Publication

Year could give insight into the effect of age, but the multivariate statistics of Publication Year only reflect positive relation in the use of innovative /w/ variants among young(er) authors. In other words, the relation between the age of authors and the innovative /v/ usage is negative. The reason is possibly explained by age-grading which caused millennial writers to change their writing behaviours (Labov, 1994; Matsuda, 2003; Tagliamonte, 2011) to conform to the katakana guidelines endorsed by the Japanese bureaucracies.

Like the other social variables mentioned above, some findings regarding the effect of genres are also incompatible with the hypothesis that traditional and formal registers (Other register) use more conservative katakana forms than the modern and informal archives such as Web register (Sakamoto, 2002; Tagliamonte, 2011). In fact, the traditional and official register of Other includes more innovative /w/ variants than other registers. For the /w/ variables, the Other register has the highest probability (94% more likely than in Book) to include innovative /w/ forms. This phenomenon may be due to the fact that the regulations or guidelines for katakana use are recommended by many Japanese divisions, but a reliable recommendation is still not established (Unger, 1984). This means that the authorities have flexibility in using various katakana forms (Daulton, 2008) which are believed to be intelligible for readers. Accordingly, this also answers why there are always multiple variants for each donor words, such as "vodka" which has six variants in the corpus. Perhaps, this also motivates the persistent effort in research and surveys regarding the katakana guidelines by NHK Broadcasting Culture Research Institute since a decade ago (NHK, 2006-2017) so that the Japanese speakers and writers do not marginalize the usage of different katakana forms.

Finally, the expectation that the use of innovative variants should be increasing over time (JTCA, 2015) is met by the patterning of innovative /w/ variants (Figure 4.6). The usage of

innovative /v/ variants, on the other hand, fails to support the hypothesis as the most recent publications (in the 2000's) show a decrease in the usage of innovative /v/ forms (Figure 4.14). I hypothesize that this decrease is influenced by the guidelines which emphasize conservative /b/ forms rather than innovative /v/ forms by the Japanese bureaucracies particularly after the two most prestigous authorities of Monbusho (1991) and Bunkacho (1991) published the katakana guidelines which became vital references for other authorities in designing their katakana guidelines (NINJAL, 2003a & 2003b & 2004 & 2006; JTCA, 2015).

In conclusion, this study has shown some surprising results with respect to the social factors. Factors such as gender have nearly no role in the use of innovative variants; younger writers have gradually disfavoured innovative variants; and the official registers do not accomplish their duty as a model for others in conserving the conservative forms recommended by the Japanese bureaucracies.

6.2 Limitations of Research

Like other studies, there were limitations in this research project. The main deficiency of this study is that for the majority of the data, we do not know the writer's age or gender (see Tables 3.4 and 3.5). Perhaps the multivariate results of gender and age were skewed by this issue. Gender and age show no effect on the use of innovative variants. In order to clarify the role of age in the innovative usage, this study applied the descriptive data of Publication Year to represent the factor of age with the notion of older publications are authored by older writers and younger writers could only possibly contribute to in the recent publications. The interpretation from Publication Year has barely shown that younger writers use more novel variants than older writers (section 5.2.3) and this is probably not as convincing as noticed from the descriptive data from Figures 4.4 and 4.12.

6.3 Future Directions

As this study used only the Chunagon corpus, this kind of research could be expanded to other databases from NINJAL[34] such as Corpus of Spontaneous Japanese (CSJ), NINJAL Web Japanese Corpus, Nagoya University Conversation Corpus, etc.

Also, instead of looking at English origin loanwords, loanwords from other donor languages are worthy of attention as well. This is because Japanese has imported an enormous amount of foreign words from various languages over time and katakanized them using conservative and possible innovative variants. In other words, other innovative katakana forms, as listed in Stanlaw (2002) mostly appear in non-English words, so the collaborative results should be interesting if extending this kind of investigation to those innovative forms. In particular, this kind of contrastive analysis may be extended to other innovative katakana forms in order to contribute more insights into the stigma of using some innovative forms.

From my point of view, extending linguistic and sociolinguistic studies to innovative katakana forms offers significant insights on the appearance of these novel katakana forms in Japanese.

[34] Of course, there are a lot of other Japanese corpora which are not constructed by NINJAL. However, at this point, corpora from NINJAL are the most establish and useful.

REFERENCES

Baker, P. (2010). *Sociolinguistics and Corpus Linguistics*. Edinburgh Sociolinguistics: Edinburgh University Press.

Bălan, C. (2015). Vowel epenthesis in Japanese loanword adaptation. *Bucharest Working Papers in Linguistics*, *17*(1), 43-68.

Barrs, K. (2011). Unlocking the encoded English vocabulary in the Japanese language. *English Today*, *27*(3), 15–20.

BKK= Bunkachō Bunkabu Kokugoka (文化庁文化部国語課) (Agency for Cultural Affairs, Culture Bureau, National Language Division). (2010). 国語に関する世論調査平成２２年度 (Opinion Poll on the National Language, 2010). Retrieved from http://www.bunka.go.jp/tokei_hakusho_shuppan/tokeichosa/kokugo_yoronchosa/pdf/h22_chosa_kekka.pdf

BKK. (2012). 国語に関する世論調査平成２４年度(Opinion Poll on the National Language, 2012). Retrieved from http://www.bunka.go.jp/tokei_hakusho_shuppan/tokeichosa/kokugo_yoronchosa/pdf/h24_chosa_kekka.pdf

Bunkacho (Agency for Cultural Affairs). (1991). 外来語の表記 (Loanword Orthography). Retrieved from http://www.bunka.go.jp/kokugo_nihongo/sisaku/joho/joho/kijun/naikaku/gairai/

Champ, N. (2014). Gairaigo in Japanese Foreign Language Learning: A Tool for Native English Speakers? *New Voices in Japanese Studies*, *6*, 117-143. Retrieved from http://newvoices.org.au/volume-6/

Chunagon corpus of NINJAL (National Institution of Japanese Language and Linguistics). Retrieved fron: https://chunagon.ninjal.ac.jp/bccwj-nt/search

Coulmas, F. (2003). *Writing systems: An introduction to their linguistic analysis.* Cambridge: Cambridge University Press.

Crawford, C. (2007). The role of loanword diffusion in changing adaptation patterns: A study of coronal stops in Japanese borrowings. *Working Papers of the Cornell Phonetics Laboratory*, *16,* 32–56.

Crawford, C. (2008). An evolutionary account of loanword-induced sound change in Japanese. University of Pennsylvania: *Working Papers in Linguistics*, *14*(1), Article 5.

Daulton, F. E. (2004). The Creation and Comprehension of English Loanwords in the Japanese Media. *Journal of Multilingual and Multicultural Development*, *25*(4), 285–296.

Daulton, F. E. (2008). *Japan's built-in lexicon of English-based loanwords*. Clevedon, UK: Multilingual Matters.

Daulton, F. E. (2015). English-based Loanwords in Japan's English Classes. Retrieved from http://repo.lib.ryukoku.ac.jp/jspui/handle/10519/6442

Dougill, J. (2008). Japan and English as an alien language. *English Today*, *24*(1), 18–22.

Fais, L., Kajikawa, S., Werker, J., Amano, S. (2005). Japanese Listeners' Perception of Phonotactic Violations. *Language and Speech*, *48*(2), 185-201.

Fotos, S. (1990). Japanese-English code switching in bilingual children. *JALT Journal*, *12*(1), 75-98.

Goldstein, D. (2011). The Use of English in Japanese Advertising. Retrieved from http://repository.cmu.edu/cgi/viewcontent.cgi?article=1120&context=hsshonors

Gordon, M. J. (2013). *Labov: A Guide for the Perplexed*. Bloomsbury Publishing.

Hobbs, V., Matsuo, A., & Payne, M. (2010). Code-switching in Japanese language classrooms: An exploratory investigation of native vs. non-native speaker teacher practice. *Linguistics and education*, *21*(1), 44-59.

Hogan, J. (2003). The social significance of English usage in Japan. *Japanese studies*, *23*, 43–58.

Igarashi, Y. (2008). *The Changing Role of Katakana in The Japanese Writing System: Processing and Pedagogical Dimensions for Native Speakers and Foreign Learners.* Saarbrucken: VDM Verlag Dr. Mueller e.K.

Inagawa, M. (2014). A re-examination of loanwords as an aid for English language learning and teaching. *International Journal of Liberal Arts and Social Science*, *4,* 13–26.

Inagawa, M. (2015). Creative and innovative uses of English in contemporary Japan: A localised taste of the global in manners posters and advertising texts. *English Today 31*(3). 11-16.

Irwin, M. (2011). *Loanwords in Japanese.* Philadelphia, NL: John Benjamins Publishing Company.

Ishiwata, T. (2001). 外来語の総合的研究 *(Studies of Loan Words)*. Tokyo: Tokyo-Dou Shuppan.

JTCA (Japan Technical Communicators Association). (2015). 外来語（カタカナ）表記ガイドライン第 3 版 (Katakana Loanword Orthography Guideline, 3rd version). Retrieved from https://www.jtca.org/standardization/katakana_guide_3_20171222.pdf

Kay, G. (1995). English loanwords in Japanese. *World Englishes*, *14*(1). 67–76.

Kochetov, A. (2008). Phonology and phonetics of loanword adaptation: Russian place names in Japanese and Korean. *Toronto Working Papers in Linguistics*, *28*, 159-174.

Kubozono, H. (2015a). Introduction. In In: Haruo Kubozono (ed.) T*he Handbook of Japanese Phonetics and Phonology*. Berlin: de Gruyter Mouton. 1-40.

Kubozono, H. (2015b). Loanword phonology. In In: Haruo Kubozono (ed.) T*he Handbook of Japanese Phonetics and Phonology*. Berlin: de Gruyter Mouton. 313-361.

Labov, W. (1972). *Sociolinguistic patterns* (No. 4). University of Pennsylvania Press.

Labov, W. (1984). Field methods of the Project on Linguistic Change and Variation. In J. Baugh & J. Sherzer(eds.), *Language in Use*. Englewood Cliffs: Prentice Hall. Intensity. GURT '84: 43-70

Labov, W. (1994). Principles of language change Volume 1: Internal factors. Oxford: Blackwell.

Labov, W. (2001). Principles of linguistic change Volume 2: Social factors. *Language in Society 29*. Oxford: Blackwell.

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y. I., & Yamada, T. (1994). Training Japanese listeners to identify English/r/and/l/. III. Long-term retention of new phonetic categories. *The Journal of the acoustical society of America*, *96*(4), 2076-2087.

Matsuda, K. (2003). Constant Rate Hypothesis, age-grading, and apparent time construct. *University of Pennsylvania Working Papers in Linguistics*, *9*(2), 11.

Monbusho (Ministry of Education of Japan). (1991). 平成 3 年 6 月28日:内閣告第 2 号:外来語の表記 (Japanese Cabinet Order No. 2: Representation of Loanwords (June 28, 1991)). Retrieved from http://www.mext.go.jp/b_menu/hakusho/nc/k19910628002/k19910628002.html

Mori, Y., & Calder, T. M. (2015). The role of motivation and learner variables in L1 and L2 vocabulary development in Japanese heritage language speakers in the United States. *Foreign Language Annals*, *48*(4), 730-754.

Nakamura, M. (2005). Developing codeswitching patterns of a Japanese/English bilingual child.

In *Proceedings of the 4th International Symposium on Bilingualism* (pp. 1679-1689).

NHK=The NHK Monthly Report on Broadcast Research and Survey. (May 2006). 外来語の発音とカタカナ表記（2） ~「wi/we/wo」などを中心に ~ (Loanword pronunciation and katakana transciption (2): "wi / we / wo" etc.). Retrieved from NHK 放送文化研究所 (NHK Broadcasting Culture Research Institute): https://www.nhk.or.jp/bunken/summary/kotoba/yougo/pdf/014.pdf

NHK. (December 2006). カタカナ表記について ~ウェブ / ウエブの表記を検討 ~ (About katakana transcription: Examination of webu/uebu). Retrieved from NHK 放送文化研究所 (NHK Broadcasting Culture Research Institute): https://www.nhk.or.jp/bunken/summary/kotoba/yougo/pdf/021.pdf

NHK. (December 2011). 外来語の表記と発音について(About loanword transcription and pronunciation). Retrieved from NHK 放送文化研究所 (NHK Broadcasting Culture Research Institute): https://www.nhk.or.jp/bunken/summary/kotoba/yougo/pdf/083.pdf

NHK. (April 2012). 外来語の表記・発音について 「ウイ・ウエ・ウオ」か「ウィ・ウェ・ウォ」か(About loanword transcription and pronunciation: /ui, ue, uo/ or /wi, we,wo/ ?). Retrieved from NHK 放送文化研究所 (NHK Broadcasting Culture Research Institute): https://www.nhk.or.jp/bunken/summary/kotoba/yougo/pdf/091.pdf

NHK. (September 2012). 外来語の発音・表記について ~〔wi〕〔we〕〔wo〕と二重母音〔ei〕~ (About loanword transcription and pronunciation: /wi, we,wo/ and diphthong /ei/). Retrieved from NHK 放送文化研究所 (NHK Broadcasting Culture Research Institute): https://www.nhk.or.jp/bunken/summary/kotoba/yougo/pdf/093.pdf

NHK. (December 2012). 外来語の発音・表記について ~ ［wei］のカタカナ表記と語末の長音 ~ (About loanword transcription and pronunciation: katakana transcription for /wei/ and long sound in word final). Retrieved from NHK 放送文化研究所 (NHK Broadcasting Culture Research Institute): https://www.nhk.or.jp/bunken/summary/kotoba/yougo/pdf/095.pdf

NHK. (February 2013). 用語の決定 ~ウオッカ，アップリケ ~ (Term decision: “vodka”,

"applique"). Retrieved from NHK 放送文化研究所 (NHK Broadcasting Culture Research Institute): https://www.nhk.or.jp/bunken/summary/kotoba/yougo/pdf/099.pdf

NHK. (May 2013). 外来語の発音・表記について ~『NHK ことばのハンドブック第 2 版』「外国語・外来語のカナ表記」細則書き換え案 ~ (About loanword transcription and pronunciation: NHK Handbook 2nd edition-revised proposal). Retrieved from NHK 放送文化研究所 (NHK Broadcasting Culture Research Institute): https://www.nhk.or.jp/bunken/summary/kotoba/yougo/pdf/103.pdf

NHK. (March 2015). 外来語の発音・表記について ヿ［wi］［we］［wo］のつく外来語の発音・表記を中心に ヿ (About loanword transcription and pronunciation: focus on transcription and pronunciation of foreign /wi/, /we/, /wo/). Retrieved from NHK 放送文化研究所 (NHK Broadcasting Culture Research Institute): https://www.nhk.or.jp/bunken/research/kotoba/pdf/20150401_5.pdf

NHK. (May 2015). 外来語の発音・表記について ~ 用語の決定 ~ (About loanword transcription and pronunciation: term decision). Retrieved from NHK 放送文化研究所 (NHK Broadcasting Culture Research Institute): https://www.nhk.or.jp/bunken/research/kotoba/pdf/20150601_5.pdf

NHK. (May 2017). 点検！外来語の伝え方 (Inspection! The way to convey loanword). Retrieved from NHK 放送文化研究所 (NHK Broadcasting Culture Research Institute): https://www.nhk.or.jp/bunken/research/kotoba/pdf/20170601_1.pdf

NHK. (June 2017). 外来語としての「アルファベット」の発音 (Pronunciation of "alphabet" in foreign words). Retrieved from NHK 放送文化研究所 (NHK Broadcasting Culture Research Institute): https://www.nhk.or.jp/bunken/research/kotoba/pdf/20170701_1.pdf

NINJAL (National Institution of Japanese Language and Linguistics). (2003a). 第 1 回「外来語」言い換え提案 (Paraphrasing Loanwords, Proposal 1), Tokyo: Kokuritsu Kokugo Kenkyoujo. Retrieved from http://pj.ninjal.ac.jp/gairaigo/Teian1/iikae_teian1.pdf

NINJAL. (2003b). 第 2 回「外来語」言い換え提案 (Paraphrasing Loanwords, Proposal 2), Tokyo: Kokuritsu Kokugo Kenkyoujo. Retrieved from

http://pj.ninjal.ac.jp/gairaigo/Teian2/iikae_teian2.pdf

NINJAL. (2004). 第３回「外来語」言い換え提案 (Paraphrasing Loanwords, Proposal 3), Tokyo: Kokuritsu Kokugo Kenkyoujo. Retrieved from http://pj.ninjal.ac.jp/gairaigo/Teian3/iikae_teian3.pdf

NINJAL. (2006). 第４回 「外来語」言い換え提案 (Paraphrasing Loanwords, Proposal 4), Tokyo: Kokuritsu Kokugo Kenkyoujo. Retrieved from http://pj.ninjal.ac.jp/gairaigo/Teian1_4/iikae_teian4.pdf

Okada, Hideo. (1999). "Japanese", in International Phonetic Association, Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet, Cambridge University Press, pp. 117–119, ISBN 978-0-52163751-0

Olah, B. (2007). English loanwords in Japanese: Effects, attitudes and usage as a means of improving spoken English ability. *Bunkyo Gakuin Daigaku Ningen-gakubu Kenkyuu Kiyo*, *9*(1). 177–188.

Paradis, C., & LaCharité, D. (2011). Loanword adaptation: From lessons learned to findings. *The Handbook of Phonological Theory, Second Edition*, 751-778.

Peperkamp, S., Vendelin, I., & Nakamura, K. (2008). On the perceptual origin of loanword adaptations: experimental evidence from Japanese. *Phonology*, *25*(1), 129-164.

Pinter, G. (2015). The emergence of new consonant contrasts. In: Haruo Kubozono (ed.) T*he Handbook of Japanese Phonetics and Phonology*. Berlin: de Gruyter Mouton. 121-166.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rebuck, M. (2002). The function of English loanwords in Japanese. *NUCB JLCC 4*(1), 53-64.

Sakamoto, M. (2002). どうする? 外来語の表記と発音--放送と外来語全国調査 (3). (How should loanword be transcribed and pronounced? -Broadcasting and Foreign Words National Survey (3)). *放送研究と調査 (Broadcasting Research and Survey)52*(10), 50-71.

Shaad, C. (1987). *A pocket guide to reading katakana*. Weatherhill, Incorporated.

Shoji, S., Shoji K. (2014). Vowel Epenthesis and Consonant Deletion in Japanese Loanwords from English. *Proceedings of the Annual Meetings on Phonology 1(1)*.

Smith, J. (2006). Loan phonology is not all perception: Evidence from Japanese loan

doublets. *Japanese/Korean Linguistics*, *14*, 63-74.

Smith, J. (2009). Source similarity in loanword adaptation: Correspondence Theory and the posited source-language representation. *Phonological argumentation: Essays on evidence and motivation. London: Equinox*, 155-177.

Stanlaw, J. (2002). "Hire" or "Fire"? Taking AD-vantage of innovations in the Japanese syllabary system. *Language Sciences*, *24*(5–6), 537–574.

Tagliamonte, S. (2011). *Variationist sociolinguistics: Change, observation, interpretation* (Vol. 40). John Wiley & Sons.

Tamaoka, K., & Miyaoka, Y. (2003). The cognitive processing of Japanese loanwords in katakana. *Japanese Psychological Research*, *45*(2), 69-79.

Trudgill, P. (2000). *Sociolinguistics: An Introduction to Language and Society*. Penguin UK.

Tsuchida, A. (1995). English loans in Japanese: Constraints in loanword phonology. *Working Papers of the Cornell Phonetics Laboratory*, *10*, 145-164.

Unger, J. M. (1984). Japanese orthography in the computer age. *Visible Language 18*,3.

Vendelin, I., & Peperkamp, S. (2006). The influence of orthography on loanword adaptations. *Lingua*, *116*(7), 996-1007.

Warner, N., & Arai, T. (2001). The role of the mora in the timing of spontaneous Japanese speech. *The Journal of the Acoustical Society of America*, *109*(3), 1144-1156.

Wong, N. F. W. (1992). *A Study of English Loanwords in Selected Malay Short Stories* (Unpublished MA thesis). Faculty of Languages and Linguistics, Kuala Lumpur: University of Malaya.

Appendix 1

The List of Primary Loanwords with /w/ and /v/ Segments

	ENGLISH	IPA(J)	IPA(E)	POSITION		ENGLISH	IPA(J)	IPA(E)	POSITION
1	vacation	ba	veɪ	Initial	**1**	window	ui	wɪ	Initial
2	valentine	ba	væ	Initial	**2**	week	ui	wi	Initial
3	variation	ba	vɛ	Initial	**3**	winter	ui	wɪ	Initial
4	version	ba	vɜ	Initial	**4**	Winnipeg	ui	wɪ	Initial
5	violin	ba	va	Initial	**5**	Wikipedia	ui	wi	Initial
6	advantage	ba	væ	Medial	**6**	whiskey	ui	wɪ	Initial
7	cover	ba	və	Medial	**7**	Halloween	ui	wi	Medial
8	advice	ba	vaɪ	Medial	**8**	sandwich	ui	wɪ	Medial
9	advance	ba	væ	Medial	**9**	wait*	ue	we	Initial
10	silver	ba	və	Medial	**10**	web	ue	wɛ	Initial
11	video	bi	vɪ	Initial	**11**	wedding	ue	wɛ	Initial
12	visual	bi	vɪ	Initial	**12**	weapon	ue	wɛ	Initial
13	vitamin	bi	vaɪ	Initial	**13**	weather	ue	wɛ	Initial
14	vision	bi	vɪ	Initial	**14**	Norway	ue	we	Medial
15	vivid	bi;bi	vɪ; və	Initial;Medial	**15**	water	uo	wɔ	Initial
16	diving	bi	vɪ	Medial	**16**	walking	uo	wɔ	Initial
17	movie	bi	vi	Medial					
18	devil	bi	və	Medial					
19	navigation	bi	və	Medial					
20	service	bi	və	Medial					
21	Vietnam	be	vi	Initial					
22	vegetable	be	vɛ	Initial					
23	venture	be	vɛ	Initial					
24	veteran	be	vɛ	Initial					
25	velvet	be;be	vɛ; və	Initial;Medial					
26	event	be	vɛ	Medial					
27	elevator	be	veɪ	Medial					
28	motivation	be	veɪ	Medial					

29	private	be	və	Medial					
30	survey	be	ve	Medial					
31	volunteer	bo	vɑ	Initial					
32	vocabulary	bo	voʊ	Initial					
33	vocal	bo	voʊ	Initial					
34	volume	bo	vɑ	Initial					
35	voice	bo	vɔ	Initial					
36	vodka	bo*	vɑ	Initial					
37	revolution	bo	və	Medial					
38	view	bju	vj	Initial					
39	active	bu	v	Final					
40	initiative	bu	v	Final					
41	stove	bu	v	Final					
42	native	bu	v	Final					
43	naive	bu	v	Final					
44	curve	bu	v	Final					

Appendix 2

The Full List of Loanwords with /w/ Variables

	LOANWORD	Variants Frequency	
		CONSERVATIVE VARIANT	INNOVATIVE VARIANT
1	Golden Week	1	0
2	Halloween	9	199
3	Norway	2	345
4	sandwich	555	87
5	waitress	126	147
6	walking	199	718
7	water	63	691
8	Waterfield	0	1
9	Waterfront	0	53
10	Waterhouse	0	42
11	Waterloo	0	18
12	weak	3	26
13	weakness	0	5
14	weapon	5	26
15	weather	9	35
16	web	39	568
17	website	25	227
18	wedding	204	174
19	week	106	379
20	weekday	14	0
21	weekend	10	27
22	weekender	2	2
23	weekly	33	84
24	weight	376	504
25	whisky	641	285
26	Wikipedia	2	117
27	window	568	1678
28	window-less	0	1
29	window shopping	0	1
30	windowpane	1	0
31	Winnipeg	1	3
32	winter	86	27
33	Winter	0	8
34	Winters	5	19
	TOTAL	**3085**	**6497**

The Full List of Loanwords with /v/ Variables

	LOANWORD	Variants Frequency	
		CONSERVATIVE VARIANT	INNOVATIVE VARIANT
1	active	367	10
2	advance	117	2
3	advanced	11	0
4	advantage	60	1
5	adventure	121	6
6	advice	3352	39
7	alternative	1	1
8	cave	0	3
9	conversion	39	0
10	cover	2360	115
11	coverage	4	0
12	covered call	8	0
13	covered warrant	3	0
14	covering	6	0
15	convert	2	0
16	curve	1038	35
17	demotivation	2	0
18	devil	99	6
19	Devil	0	1
20	Devil May Cry	2	0
21	Devil Summoner	1	0
22	deviled egg	2	0
23	Devils	0	3
24	disadvantage	1	0
25	discover	11	0
26	diving	407	1
27	division	21	10
28	elevator	1402	41
29	event	4514	33
30	full-covered	1	0
31	hardcover	47	0
32	imaginative	1	0
33	individual	0	11
34	initiative	279	36
35	interview	1821	84
36	interviewer	4	1
37	invoice	30	1
38	motivation	262	5
39	movie	369	6
40	naive	88	24
41	native	198	105
42	navigation	95	1
43	preview	442	0

44	private	868	11
45	reactive	7	0
46	recover	4	0
47	review	443	74
48	revision	9	0
49	revolution	53	2
50	self-service	48	4
51	service	16344	139
52	serviceman	33	0
53	silver	1201	41
54	skydiving	31	0
55	softcover	5	0
56	stove	517	32
57	surveillance	89	1
58	survey	39	4
59	television	390	3
60	uncover	1	0
61	vacation	23	2
62	Valentine	459	40
63	variation	528	59
64	vegetable	55	4
65	velvet	154	28
66	venture	879	48
67	version	1323	107
68	veteran	856	12
69	video	4404	59
70	videocassette	17	0
71	videodisk	7	0
72	videotape	234	8
73	videotex	50	0
74	Vietnam	1998	189
75	view	645	41
76	viewer	65	0
77	violin	293	417
78	vision	1119	128
79	Vision	1	0
80	visual	232	56
81	vitamin	2051	3
82	vivid	132	36
83	vocabulary	44	3
84	vocal	315	171
85	vodka	150	126
86	voice	242	53
87	volume	986	24
88	volunteer	3885	27
	TOTAL	**58813**	**2533**

Appendix 3

IPA Symbols for /w/ and /v/ Variables

(R could not read all symbols of IPA vowels, therefore some of them are replaced by "readable" symbols so that R could read the spreadsheet in Microsoft Excel for statistical analysis.)

	IPA	**Symbol in Excel**
1	ɪ	I
2	ʊ	u
3	ə	uh
4	ɛ	E
5	ɜ	er
6	ɔ	O
7	æ	ae
8	ɑ	a
9	ɑɪ	aj
10	eɪ	ej

	/v/ variables			**/w/ variables**	
	IPA	**Symbol in Excel**		**IPA**	**Symbol in Excel**
1	v	v	1	wɪ	wI
2	vɑ	va	2	wɛ	wE
3	væ	vae	3	wɔ	wO
4	vaɪ	vaj	4	wi	wi
5	ve	ve	5	we	we
6	vɛ	vE			
7	veɪ	vej			
8	vɜ	ver			
9	vi	vi			
10	vɪ	vI			
11	vj	vj			
12	vɔ	vO			
13	voʊ	vou			
14	və	vuh			

Appendix 4

The IPA and R-readable Symbols for /w/ and /v/ Segments

	/V/ LW	IPA	Symbol			/W/ LW	IPA	Symbol
1	active	v	v		1	Golden Week	wi	wi
2	advance	væ	vae		2	Halloween	wi	wi
3	advanced	væ	vae		3	Norway	we	we
4	advantage	væ	vae		4	sandwich	wɪ	wI
5	adventure	vɛ	vE		5	waitress	we	we
6	advice	vaɪ	vaj		6	walking	wɔ	wO
7	alternative	v	v		7	water	wɔ	wO
8	cave	v	v		8	Waterfield	wɔ	wO
9	conversion	vɜ	ver		9	waterfront	wɔ	wO
10	cover	və	vuh		10	Waterhouse	wɔ	wO
11	coverage	və	vuh		11	Waterloo	wɔ	wO
12	covered call	və	vuh		12	weak	wi	wi
13	covered warrant	və	vuh		13	weakness	wi	wi
14	covering	və	vuh		14	weapon	wɛ	wE
15	covert	və	vuh		15	weather	wɛ	wE
16	curve	v	v		16	web	wɛ	wE
17	demotivation	veɪ	vej		17	website	wɛ	wE
18	devil	və	vuh		18	wedding	wɛ	wE
19	Devil	və	vuh		19	week	wi	wi
20	Devil May Cry	və	vuh		20	weekday	wi	wi
21	Devil Summoner	və	vuh		21	weekend	wi	wi
22	deviled egg	və	vuh		22	weekender	wi	wi
23	Devils	və	vuh		23	weekly	wi	wi
24	disadvantage	væ	vae		24	weight	we	we
25	discover	və	vuh		25	whisky	wɪ	wI
26	diving	vɪ	vI		26	Wikipedia	wi	wi
27	division	vɪ	vI		27	window	wɪ	wI
28	elevator	veɪ	vej		28	window-less	wɪ	wI
29	event	vɛ	vE		29	window shopping	wɪ	wI
30	full-covered	və	vuh		30	windowpane	wɪ	wI
31	hardcover	və	vuh		31	Winnipeg	wɪ	wI
32	imaginative	v	v		32	winter	wɪ	wI
33	individual	vɪ	vI		33	Winter	wɪ	wI
34	initiative	v	v		34	Winters	wɪ	wI
35	interview	vj	vj					
36	interviewer	vj	vj					
37	invoice	vɔ	vO					
38	motivation	veɪ	vej					

39	movie	vi	vi					
40	naive	v	v					
41	native	v	v					
42	navigation	və	vuh					
43	preview	vj	vj					
44	private	və	vuh					
45	reactive	v	v					
46	recover	və	vuh					
47	review	vj	vj					
48	revision	vɪ	vI					
49	revolution	və	vuh					
50	self-service	və	vuh					
51	service	və	vuh					
52	serviceman	və	vuh					
53	silver	və	vuh					
54	skydiving	vɪ	vI					
55	softcover	və	vuh					
56	stove	v	v					
57	surveillance	ve	ve					
58	survey	ve	ve					
59	television	vɪ	vI					
60	uncover	və	vuh					
61	vacation	veɪ	vej					
62	Valentine	væ	vae					
63	variation	vɛ	vE					
64	vegetable	vɛ	vE					
65	velvet	vɛ; və	vE; vuh					
66	venture	vɛ	vE					
67	version	vɜ	ver					
68	veteran	vɛ	vE					
69	video	vɪ	vI					
70	videocassette	vɪ	vI					
71	videodisk	vɪ	vI					
72	videotape	vɪ	vI					
73	videotex	vɪ	vI					
74	Viet Nam	vi	vi					
75	view	vj	vj					
76	viewer	vj	vj					
77	violin	vaɪ	vaj					
78	vision	vɪ	vI					
79	Vision	vɪ	vI					

80	visual	vɪ	vI					
81	vitamin	vaɪ	vaj					
82	vivid	vɪ ; və	vI; vuh					
83	vocabulary	voʊ	vou					
84	vocal	voʊ	vou					
85	vodka	vɑ	va					
86	voice	vɔ	vO					
87	volume	vɑ	va					
88	volunteer	vɑ	va					

Appendix 5

The Top Ten "Oldest" Loanwords Found in the Corpus

SAMPLE ID	ENGLISH	KATAKANA	ENGLISH SEGNENT	SEGMENT	FORM	POSITION	GENDER	YOB	SUB-REGISTER	REGISTER	PUBLISHED YEAR
OB0X_00016	initiative	inishiatibu	v	bu	conservative	final	N	0	Sbestseller	Book	1975
OB0X_00009	stove	suto:bu	v	bu	conservative	final	M	1910	Sbestseller	Book	1975
OB0X_00026	whisky	uisuki:	wI	ui	conservative	initial	M	1930	Sbestseller	Book	1975
OB0X_00030	vacation	vake:shon	vej	va	innovative	initial	M	1920	Sbestseller	Book	1975
OB0X_00030	Viet Nam	vetonamu	vi	ve	innovative	initial	M	1920	Sbestseller	Book	1975
OB0X_00030	sandwich	sandowicchi	wI	wi	innovative	medial	M	1920	Sbestseller	Book	1975
OB0X_00016	week	wi:ku	wi	wi	innovative	initial	M	0	Sbestseller	Book	1975
OB0X_00030	weekend	wi:kuendo	wi	wi	innovative	initial	M	1920	Sbestseller	Book	1975
OB0X_00029	window	windo:	wI	wi	innovative	initial	M	1920	Sbestseller	Book	1975
OB0X_00030	water	wo:ta:	wO	wo	innovative	initial	M	1920	Sbestseller	Book	1975

Appendix 6

The Appearance of Historical Form of /i/ for “Sandwich” in the Corpus from the 1970’s to the 2000’s (Not Full-listed)

SAMPLE ID	ENGLISH	KATAKANA	ENGLISH SEG	SEGMENT	FORM	SUB-REGISTER	REGISTER	PUBLISHED YEAR
OB1X_00084	sandwich	sandoicchi	wI	I	conservative	Sbestseller	Book	1977
OB1X_00145	sandwich	sandoicchi	wI	I	conservative	Sbestseller	Book	1978
OM25_00008	sandwich	sandoicchi	wI	I	conservative	Sdietrecord	Other	1981
OB2X_00308	sandwich	sandoicchi	wI	I	conservative	Sbestseller	Book	1982
LBa7_00001	sandwich	sandoicchi	wI	I	conservative	Lbook	Book	1986
OM34_00001	sandwich	sandoicchi	wI	I	conservative	Sdietrecord	Other	1986
LBb9_00060	sandwich	sandoicchi	wI	I	conservative	Lbook	Book	1987
OB3X_00202	sandwich	sandoicchi	wI	I	conservative	Sbestseller	Book	1988
LBd6_00005	sandwich	sandoicchi	wI	I	conservative	Lbook	Book	1989
LBe9_00006	sandwich	sandoicchi	wI	I	conservative	Lbook	Book	1990
LBf9_00082	sandwich	sandoicchi	wI	I	conservative	Lbook	Book	1991
LBg9_00054	sandwich	sandoicchi	wI	I	conservative	Lbook	Book	1992
LBh2_00029	sandwich	sandoicchi	wI	I	conservative	Lbook	Book	1993
LBi9_00241	sandwich	sandoicchi	wI	I	conservative	Lbook	Book	1994
LBj9_00186	sandwich	sandoicchi	wI	I	conservative	Lbook	Book	1995
LBk2_00077	sandwich	sandoicchi	wI	I	conservative	Lbook	Book	1996
LBl9_00107	sandwich	sandoicchi	wI	I	conservative	Lbook	Book	1997
LBm9_00258	sandwich	sandoicchi	wI	I	conservative	Lbook	Book	1998
LBn3_00007	sandwich	sandoicchi	wI	I	conservative	Lbook	Book	1999
LBo8_00007	sandwich	sandoicchi	wI	I	conservative	Lbook	Book	2000
PB12_00004	sandwich	sandoicchi	wI	I	conservative	Pbook	Book	2001
PM11_00600	sandwich	sandoicchi	wI	I	conservative	Pmagazine	Publication	2001
PB26_00042	sandwich	sandoicchi	wI	I	conservative	Pbook	Book	2002
PM21_01297	sandwich	sandoicchi	wI	I	conservative	Pmagazine	Publication	2002
PB39_00648	sandwich	sandoicchi	wI	I	conservative	Pbook	Book	2003
PM31_00100	sandwich	sandoicchi	wI	I	conservative	Pmagazine	Publication	2003
OB6X_00181	sandwich	sandoicchi	wI	I	conservative	Sbestseller	Book	2004
OW6X_00475	sandwich	sandoicchi	wI	I	conservative	Swhitepaper	Other	2004
PN4k_00009	sandwich	sandoicchi	wI	I	conservative	Pnewspaper	Publication	2004
PB52_00128	sandwich	sandoicchi	wI	I	conservative	Pbook	Book	2005
PM51_00636	sandwich	sandoicchi	wI	I	conservative	Pmagazine	Publication	2005
OC14_11433	sandwich	sandoicchi	wI	I	conservative	Syahooanswers	Web	2005
OT63_00039	sandwich	sandoicchi	wI	I	conservative	Stextbook	Book	2006
OP21_00003	sandwich	sandoicchi	wI	I	conservative	Spublicpaper	Publication	2008
OY15_21004	sandwich	sandoicchi	wI	I	conservative	Syahooblogs	Web	2008

Appendix 7

The Conservative Form /ui/ for “Sandwich” which Appeared only in the 2000’s on Web

SAMPLE ID	ENGLISH	KATAKANA	ENGLISH SEGNENT	SEGMENT	FORM				SUB-REGISTER	REGISTER	PUBLISHED YEAR
OC14_08445	sandwich	sandouicchi	wI	ui	conservative	medial	N	0	Syahooanswers	Web	2005
OC14_08445	sandwich	sandouicchi	wI	ui	conservative	medial	N	0	Syahooanswers	Web	2005
OC14_08445	sandwich	sandouicchi	wI	ui	conservative	medial	N	0	Syahooanswers	Web	2005
OY08_01165	sandwich	sandouicchi	wI	ui	conservative	medial	N	0	Syahooblogs	Web	2008
OY15_16778	sandwich	sandouicchi	wI	ui	conservative	medial	N	0	Syahooblogs	Web	2008

Appendix 8

The Appearance of Intermediate Form of /uwi/ for “Whisky” in the Corpus from the 1970’s to the 2000’s

SAMPLE ID	ENGLISH	KATAKANA	ENGLISH SEGNENT	SEGMENT	FORM	POSITION	GENDER	YOB	SUB-REGISTER	REGISTER	PUBLISHED YEAR
OM11_00008	whisky	uwisuki:	wI	uwi	conservative	initial	N	0	Sdietrecord	Other	1978
LBf9_00042	whisky	uwisukii	wI	uwi	conservative	initial	M	1910	Lbook	Book	1991
LBl5_00044	whisky	uwisuki:	wI	uwi	conservative	initial	N	0	Lbook	Book	1997
LBl5_00044	whisky	uwisuki:	wI	uwi	conservative	initial	N	0	Lbook	Book	1997
LBl5_00044	whisky	uwisuki:	wI	uwi	conservative	initial	N	0	Lbook	Book	1997
LBl5_00044	whisky	uwisuki:	wI	uwi	conservative	initial	N	0	Lbook	Book	1997
LBl5_00044	whisky	uwisuki:	wI	uwi	conservative	initial	N	0	Lbook	Book	1997
LBn6_00017	whisky	uwisuki:	wI	uwi	conservative	initial	M	1940	Lbook	Book	1999
LBn6_00017	whisky	uwisuki:	wI	uwi	conservative	initial	M	1940	Lbook	Book	1999
LBr5_00050	whisky	uwisuki:	wI	uwi	conservative	initial	M	1940	Lbook	Book	2003

Appendix 9

The Appearance of Historical Form of /ui/ for “Whisky” in the Corpus from the 1970’s to the 2000’s (Not Full-listed)

SAMPLE ID	ENGLISH	KATAKANA	ENGLISH SEGNENT	SEGMENT	FORM	SUB-REGISTER	REGISTER	PUBLISHED YEAR
OB0X_00026	whisky	uisuki:	wI	ui	conservative	Sbestseller	Book	1975
OB1X_00023	whisky	uisuki:	wI	ui	conservative	Sbestseller	Book	1976
OB1X_00173	whisky	uisuki:	wI	ui	conservative	Sbestseller	Book	1977
OB1X_00128	whisky	uisuki:	wI	ui	conservative	Sbestseller	Book	1978
OW1X_00297	whisky	uisuki:	wI	ui	conservative	Swhitepaper	Other	1978
OB1X_00041	whisky	uisuki:	wI	ui	conservative	Sbestseller	Book	1980
OB2X_00001	whisky	uisuki:	wI	ui	conservative	Sbestseller	Book	1981
OB2X_00017	whisky	uisuki:	wI	ui	conservative	Sbestseller	Book	1982
OB2X_00139	whisky	uisuki:	wI	ui	conservative	Sbestseller	Book	1983
OW2X_00180	whisky	uisuki:	wI	ui	conservative	Swhitepaper	Other	1984
OB2X_00116	whisky	uisuki:	wI	ui	conservative	Sbestseller	Book	1985
OB3X_00013	whisky	uisuki:	wI	ui	conservative	Sbestseller	Book	1986
LBb7_00011	whisky	uisuki:	wI	ui	conservative	Lbook	Book	1987
OM32_00002	whisky	uisuki:	wI	ui	conservative	Sdietrecord	Other	1987
LBc9_00144	whisky	uisuki:	wI	ui	conservative	Lbook	Book	1988
LBd9_00100	whisky	uisuki:	wI	ui	conservative	Lbook	Book	1989
LBe5_00018	whisky	uisuki:	wI	ui	conservative	Lbook	Book	1990
LBf9_00224	whisky	uisuki:	wI	ui	conservative	Lbook	Book	1991
LBg9_00001	whisky	uisuki:	wI	ui	conservative	Lbook	Book	1992
LBh9_00127	whisky	uisuki:	wI	ui	conservative	Lbook	Book	1993
OV2X_00036	whisky	uisuki:	wI	ui	conservative	Sverse	Other	1993
LBi9_00236	whisky	uisuki:	wI	ui	conservative	Lbook	Book	1994
LBj9_00060	whisky	uisuki:	wI	ui	conservative	Lbook	Book	1995
OB5X_00067	whisky	uisuki:	wI	ui	conservative	Sbestseller	Book	1996
OB5X_00248	whisky	uisuki:	wI	ui	conservative	Sbestseller	Book	1997
LBm9_00176	whisky	uisuki:	wI	ui	conservative	Lbook	Book	1998
OB5X_00179	whisky	uisuki:	wI	ui	conservative	Sbestseller	Book	1999
LBo9_00240	whisky	uisuki:	wI	ui	conservative	Lbook	Book	2000
PB19_00085	whisky	uisuki:	wI	ui	conservative	Pbook	Book	2001
PM11_00227	whisky	uisuki:	wI	ui	conservative	Pmagazine	Publication	2001
PB29_00575	whisky	uisuki:	wI	ui	conservative	Pbook	Book	2002
PM21_01335	whisky	uisuki:	wI	ui	conservative	Pmagazine	Publication	2002
PB35_00016	whisky	uisuki:	wI	ui	conservative	Pbook	Book	2003
PM32_00013	whisky	uisuki:	wI	ui	conservative	Pmagazine	Publication	2003
PB49_00019	whisky	uisuki:	wI	ui	conservative	Pbook	Book	2004
PN4g_00008	whisky	uisuki:	wI	ui	conservative	Pnewspaper	Publication	2004
PB59_00358	whisky	uisuki:	wI	ui	conservative	Pbook	Book	2005
PN5g_00005	whisky	uisuki:	wI	ui	conservative	Pnewspaper	Publication	2005
OC14_06893	whisky	uisuki:	wI	ui	conservative	Syahooanswers	Web	2005
OP77_00001	whisky	uisuki:	wI	ui	conservative	Spublicpaper	Publication	2008
OY15_12905	whisky	uisuki:	wI	ui	conservative	Syahooblogs	Web	2008

Appendix A

Hiragana Charts

Vowel	/a/	/i/	/u/	/e/	/o/		/a/	/i/	/u/	/e/	/o/
Consonant											
-	あ	い	う	え	お						
	a	i	u	e	o						
/k/	か	き	く	け	こ		が	ぎ	ぐ	げ	ご
	ka	ki	ku	ke	ko		ga	gi	gu	ge	go
/s/	さ	し	す	せ	そ		ざ	じ	ず	ぜ	ぞ
	sa	ʃi	su	se	so		za	zi	zu	ze	zo
/t/	た	ち	つ	て	と		だ	ぢ	づ	で	ど
	ta	ʧi	tsu	te	to		da	ʤi	dzu	de	do
/n/	な	に	ぬ	ね	の						
	na	ni	nu	ne	no						
/h/	は	ひ	ふ	へ	ほ		ば	び	ぶ	べ	ぼ
	ha	hi	hu	he	ho		ba	bi	bu	be	bo
							ぱ	ぴ	ぷ	ぺ	ぽ
							pa	pi	pu	pe	po
/m/	ま	み	む	め	も						
	ma	mi	mu	me	mo						
/r/	ら	り	る	れ	ろ						
	ra	ri	ru	re	ro						
/j/	や		ゆ		よ						
	ja		ju		jo						
/w/	わ				を						
	wa				wo						
	ん										
	n										

Chart A.1: Hiragana characters: Basic characters and the voiced counterparts for /k/, /s/, /t/ and /h/ rows.

Glides	/ja/		/ju/		/jo/		/ja/		/ju/		/jo/
Consonant followed by /i/											
/ki/	きゃ		きゅ		きょ		ぎゃ		ぎゅ		ぎょ
	kja		kju		kjo		gja		gju		gjo
/ʃi/	しゃ		しゅ		しょ		じゃ		じゅ		じょ
	ʃa		ʃu		ʃo		ʤa		ʤu		ʤo
/ʧi/	ちゃ		ちゅ		ちょ						
	ʧa		ʧu		ʧo						
/ni/	にゃ		にゅ		にょ						
	nja		nju		njo						
/hi/	ひゃ		ひゅ		ひょ		びゃ		びゅ		びょ
	ça		çu		ço		bja		bju		bjo
							ぴゃ		ぴゅ		ぴょ
							pja		pju		pjo
/mi/	みゃ		みゅ		みょ						
	mja		mju		mjo						
/ri/	りゃ		りゅ		りょ						
	rja		rju		rjo						

Chart A.2: Hiragana characters: Contracted characters and the voiced counterparts for /k/, /s/, and /h/ rows.

Appendix B

Katakana Charts

Vowel	/a/	/i/	/u/	/e/	/o/		/a/	/i/	/u/	/e/	/o/
Consonant											
-	ア	イ	ウ	エ	オ						
	a	i	u	e	o						
/k/	カ	キ	ク	ケ	コ		ガ	ギ	グ	ゲ	ゴ
	ka	ki	ku	ke	ko		ga	gi	gu	ge	go
/s/	サ	シ	ス	セ	ソ		ザ	ジ	ズ	ゼ	ゾ
	sa	ʃi	su	se	so		za	zi	zu	ze	zo
/t/	タ	チ	ツ	テ	ト		ダ	ヂ	ヅ	デ	ド
	ta	ʧi	tsu	te	to		da	ʤi	dzu	de	do
/n/	ナ	ニ	ヌ	ネ	ノ						
	na	ni	nu	ne	no						
/h/	ハ	ヒ	フ	ヘ	ホ		バ	ビ	ブ	ベ	ボ
	ha	hi	hu	he	ho		ba	bi	bu	be	bo
							パ	ピ	プ	ペ	ポ
							pa	pi	pu	pe	po
/m/	マ	ミ	ム	メ	モ						
	ma	mi	mu	me	mo						
/r/	ラ	リ	ル	レ	ロ						
	ra	ri	ru	re	ro						
/j/	ヤ		ユ		ヨ						
	ja		ju		jo						
/w/	ワ				ヲ						
	wa				wo						
	ン										
	n										

Chart B.1: Katakana characters: Basic characters and the voiced counterparts for /k/, /s/, /t/ and /h/ rows.

Glides	/ja/		/ju/		/jo/		/ja/		/ju/		/jo/
Consonant followed by /i/											
/ki/	キャ		キュ		キョ		ギャ		ギュ		ギョ
	kja		kju		kjo		gja		gju		gjo
/ʃi/	シャ		シュ		ショ		ジャ		ジュ		ジョ
	ʃa		ʃu		ʃo		ʤa		ʤu		ʤo
/ʧi/	チャ		チュ		チョ						
	ʧa		ʧu		ʧo						
/ni/	ニャ		ニュ		ニョ						
	nja		nju		njo						
/hi/	ヒャ		ヒュ		ヒョ		ビャ		ビュ		ビョ
	ça		çu		ço		bja		bju		bjo
							ピャ		ピュ		ピョ
							pja		pju		pjo
/mi/	ミャ		ミュ		ミョ						
	mja		mju		mjo						
/ri/	リャ		リュ		リョ						
	rja		rju		rjo						

Chart B.2: Katakana characters: Contracted characters and the voiced counterparts for /k/, /s/, and /h/ rows.

Appendix C

Innovative Katakana Chart

1950's	**1960's**	**1970's**	**1980's**	**Early 1990's**	**Later 1990's**	**2000's**
ジェ	シェ	ズィ	ツァ	イェ	グィ	スィ
ʤe	ʃe	zi	tsa	je	gwi	si
チェ	**ヴァ**	トゥ	ツィ	クヮ	グェ	フゥ
ʧe	**va**	tu	tsi	kwa	gwe	hu
ティ	**ヴィ**	ドゥ	ツェ	クィ	グォ	
ti	**vi**	du	tse	kwi	gwo	
ディ	**ヴ**	テュ	ツォ	クェ	スァ	
di	**vu**	tju	tso	kwe	swa	
ファ	**ヴェ**	デュ	ヴュ	クォ	スィ	
fa	**ve**	dju	vju	kwo	swi	
フィ	**ヴォ**	フュ		グヮ	スェ	
fi	**vo**	fju		gwa	swe	
フォ	**ウォ**					
fo	**wo**					
ウィ						
wi						
ウェ						
we						

Chart C: Innovative katakana lists based on their occurrence times from dictionary data *(sourced from Stanlaw, 2002, p. 569)*