

**A Canadian Perspective on  
Japanese-English Language Contact**

Yukiko Yoshizumi

Thesis submitted to the  
Faculty of Graduate and Postdoctoral Studies  
In partial fulfillment of the requirements  
For the PhD degree in Linguistics

Department of Linguistics  
Faculty of Arts  
University of Ottawa

## ABSTRACT

This dissertation investigates the linguistic outcomes of Japanese-English language contact in Canada. Adopting a sociolinguistic variationist framework (Labov 1966; Sankoff & Labov 1985), the main objective is to determine whether or not Japanese spoken in Canada (hereafter, heritage Japanese) is showing structural change due to prolonged contact with English. The study is based on naturalistic speech data collected from 16 Japanese-English bilingual speakers in Canada. A key component of this dissertation is the use of a comparative sociolinguistic framework (Poplack and Tagliamonte 2001; Tagliamonte 2002) to assess structural affinities between heritage Japanese and the homeland Japanese benchmark variety. Speech patterns in heritage Japanese are systematically compared with patterns found in a commensurate monolingual benchmark variety of Japanese with regard to three linguistic *variables*, which are considered to be vulnerable to contact-induced language change (i.e. Bullock 2004, Sorace 2011). In terms of the first variable analyzed, *variable realization of subject pronouns*, it was found that the underlying grammar in heritage Japanese is shared by the homeland benchmark variety, showing that the variable is conditioned by the factor groups of *subject continuity* (i.e. switch reference) and *grammatical person*; the null variant is favoured by the same subject referent and the second person pronoun. Second, with regard to *variable case marking on subject nouns* and *variable case marking on direct object nouns*, it was found that the same underlying grammar is shared for case marking. For example, the constraint hierarchies in heritage Japanese were identical with those in the homeland variety for focus particles, with presence of a focus particle favouring null marking consistently for all types of nouns (i.e. English-origin nouns and Japanese nouns in heritage Japanese, and Japanese nouns and loanwords in homeland Japanese). The constraint hierarchies (and direction of the effect) for the other significant factor groups of verbal adjacency and sentence-final particle were identical between heritage Japanese and the homeland variety, with the exception of a reversed direction of effect for loanword subject nouns in heritage Japanese for the non-significant factor group of verbal adjacency, and a neutralized effect for Japanese nouns in heritage Japanese and loanwords in homeland Japanese when these nouns are located in direct object position. Considered in the aggregate, constraint hierarchies were found to exhibit a number of parallels across comparison varieties. This finding bolsters the general conclusion that there is little evidence indicating that extensive contact with English has had any discernible impact on structural patterns in these sectors of the heritage grammar. Furthermore, it was shown that no social factor group (i.e. length of stay in Canada) has an appreciable effect on heritage Japanese. Summarizing, the multiple lines of evidence emerging from the empirical quantitative analyses of the variables targeted in this dissertation converge in indicating that heritage Japanese, as spoken in Canada, broadly shares the same underlying grammar as homeland Japanese. Structural affinities in variable patterning shared by heritage and homeland varieties reveal little compelling evidence indicating that heritage Japanese exhibits structural change due to contact with English.

## ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere and deepest gratitude to my supervisor, Professor Stephen Levey, for the continuous support, guidance, enthusiasm, motivation, patience, and most importantly his immense knowledge. I feel very fortunate that I was able to work with him; he cared about my work and responded to my questions and queries promptly throughout the research and writing of this dissertation. His guidance made this journey successful. For everything you have done for me, I thank you, Stephen.

I would like to acknowledge the members of the dissertation committee: Professor Naomi Nagy (University of Toronto), Professor Juana Muñoz-Liceras (University of Ottawa), Professor Laura Sabourin (University of Ottawa), and Professor Junko Shimoyama (McGill University). I am very grateful to Professor Nagy for becoming a part of the committee since her research project on heritage languages in Toronto was a great inspiration for this dissertation. I would like to thank Professor Muñoz-Liceras and Professor Sabourin for their thoroughly informative and helpful comments. I am also very grateful to Professor Shimoyama for being a part of the committee and providing me with valuable comments, particularly from the perspective of Japanese linguistics. Their insightful feedback gives me a hopeful direction for future research.

I am indebted to Professor Shana Poplack for providing me with a solid foundation on the study of language variation and change. I am very grateful to Professor Robert Truswell for providing me with references on the theoretical components. I would also like to thank all my informants for participating in the interviews and sharing their stories with me, and Professor Keven Heffernan for granting me access to the *Corpus of Kwansai Spoken Japanese* (Heffernan 2012).

All my linguistic friends deserve my gratitude. First, many thanks to my sociolinguistic fellows: Mélissa Chiasson-Léger, Nicté Fuller Medina, Joseph Roy, Allison Leales, Laura Kastronic, Ekab Al-Shawashreh, and Nathalie Dion for helping me to survive this long academic road. Special thanks to my linguistic friends for their friendship and support and for the pleasant memories: Keren Tonciulescu, Sabine Huynh, Carmen LeBlanc, Galina Dukova-Zheleva, Dana Geber, Reza Falahati, Ewelina Frackowiak, and Mutsumi Oi. I also wish to express special thanks to my original officemates with whom I shared a lot of fun times: Marie-Claude Séguin, Keiko Kaku, and Stephanie Chen. Time with them inside and outside our office, room 432, is full of good memories for me still now.

I am very thankful to Takako Sato and the late Professor Vasilis Vourkoutiotis, as well as to Naoko Ikeda and Professor Jérémie Séror for their generous support and continuous encouragement in various ways during my academic years. Takako, in particular, always created a family atmosphere for me where I really felt warm and comfortable, regardless of being away from Japan.

I would like to express my heartfelt thanks to Ewa Trębacz and Waław Trębacz for their encouragement and warmhearted personalities, which gave me a relaxing and comfortable feeling, especially when I needed to be away from studying.

I would like to thank Michał Trębacz for always being patient with me, listening to me, and influencing me in various ways. I cannot list everything he did for me, but what I can definitely say is that he was always there for me, stretching out his hand toward me so that I could continue to go forward toward the goal. I could not have done this at all without him.

Finally, I would like to express special thanks to my wonderful family, without whom I could not have gotten through this long and arduous way and achieve my PhD goal. My sister, Chisako, always cheered me up while my brother, Tetsuya, gave me energy. Most importantly, my parents, Teiko and Tomoyuki Yoshizumi, have always been there for me and believed in me on this long journey. Their strong support and love made this happen. I would like to dedicate this dissertation to my parents.

# TABLE OF CONTENTS

ABSTRACT .....	ii
ACKNOWLEDGEMENTS.....	iii
LIST OF TABLES .....	ix
LIST OF FIGURES .....	xii
LIST OF ABBREVIATIONS.....	xiii

## CHAPTER 1 INTRODUCTION

1.1 Background .....	1
1.2 Overview of the Present Study .....	4
1.3 Contribution of Japanese-English Data to the Study of Contact-Induced Change .....	7
1.4 Typological Differences between Japanese and English .....	8
1.5 Goals.....	14
1.5.1 Research Questions .....	14
1.5.2 Variable Realization of Subject Pronouns.....	15
1.5.3 Variable Case Marking on Nouns in Subject and Direct Object Positions .....	16
1.6 Theoretical Framework .....	18
1.7 Overview of the Dissertation.....	19

## CHAPTER 2 LITERATURE REVIEW

2.1 Outcomes of Contact-Induced Change .....	21
2.2 Distinguishing Borrowing and Codeswitching .....	25
2.3 Language Contact between Japanese and English .....	29
2.3.1 Japanese-English Contact in Monolingual Settings .....	29
2.3.2 Japanese-English Contact in Bilingual Settings.....	31

## CHAPTER 3 METHODOLOGY

3.1 Variationist Framework.....	35
3.1.1 The Primacy of Vernacular Data.....	36
3.1.2 The Principle of Accountability.....	37
3.1.3 Quantitative Analysis.....	37
3.2 The Comparative Method in Language Contact Studies.....	39
3.3 Japanese as a Heritage Language.....	41
3.4 Data.....	42
3.4.1 The Japanese-English Bilingual Community in Ontario (Toronto and Ottawa).....	42
3.4.2 Data and Speaker Sample.....	45
3.5 Procedure.....	54

## CHAPTER 4 VARIABLE REALIZATION OF SUBJECT PRONOUNS

4.1 Introduction.....	56
4.2 Null Subject Languages.....	58
4.2.1 Null Subject Languages in the Earlier Studies.....	58
4.2.2 Radical Pro-drop Languages.....	60
4.2.3 Typological Approach to Null Subject Languages.....	61
4.3 English as a Non-Pro-Drop Language.....	63
4.4 Previous Studies on Null Subjects in Japanese.....	65
4.5 Null Subjects in Contact Varieties.....	72
4.6 Variable Context.....	75
4.6.1 Envelope of Variation.....	75
4.6.2 Extraction Procedures and Exclusion.....	75
4.6.2.1 Extraction.....	75
4.6.2.2 Exclusions.....	77
4.7 Hypothesis.....	80
4.8 Factor Groups.....	81
4.8.1 Social Factors.....	81
4.8.2 Linguistic Factors.....	82
4.8.2.1 Grammatical Person.....	82

4.8.2.2 Subject Continuity .....	84
4.8.2.3 Previous Realization (Priming).....	86
4.8.2.4 Clause Type .....	87
4.8.2.5 Verbal Transitivity .....	88
4.9 Results .....	89
4.9.1 Overall Distribution.....	89
4.9.2 Multivariate Analysis .....	91
4.10 Discussion .....	96

## CHAPTER 5 VARIABLE CASE MARKING ON NOUNS IN SUBJECT AND DIRECT OBJECT POSITION

5.1 Introduction .....	98
5.2 Case Marking in Japanese .....	100
5.2.1 Theoretical Approach .....	100
5.2.2 Empirical Studies.....	104
5.3 Case Marking in Contact Varieties .....	107
5.4 Variable Context .....	109
5.4.1 Envelope of Variation.....	109
5.4.2 Extraction Procedures and Exclusions .....	110
5.4.2.1 Extraction .....	110
5.4.2.2 Exclusions.....	112
5.4.2.2.1 Categorical Environments.....	112
5.4.2.2.2 Proper Nouns .....	113
5.4.2.2.3 Light Verb Construction .....	113
5.5 Hypothesis.....	121
5.6 Factor Groups .....	121
5.6.1 Social Factors .....	122
5.6.2 Linguistic Factors .....	123
5.6.2.1 Verbal Adjacency.....	123
5.6.2.2 Presence of Focus Particle .....	124
5.6.2.3 Presence of Sentence-final Particle.....	125
5.6.2.4 Word Order .....	128

5.6.2.5 Clause Type .....	130
5.7 Results .....	131
5.7.1 Overall Distribution.....	131
5.7.2 Multivariate Analysis .....	135
5.7.2.1 Variable Case Marking on Nouns in Subject Position.....	136
5.7.2.2 Variable Case Marking on Nouns in Direct Object Position.....	144
5.8 Discussion.....	149
CHAPTER 6 CONCLUSION .....	152
BIBLIOGRAPHY .....	162

## LIST OF TABLES

### CHAPTER 1 INTRODUCTION

Table 1.1: Summary of dissimilarities between Japanese and English .....	13
--	----

### CHAPTER 3 METHODOLOGY

Table 3.1 Demographic profile of heritage Japanese .....	43
Table 3.2 Information about heritage Japanese speakers in Toronto in the <i>CJECC</i> .....	47
Table 3.3 Information about heritage Japanese speakers in Ottawa in the <i>CJECC</i> .....	48
Table 3.4 Composition of heritage Japanese speakers in Canada according to years of residence in Canada with more detailed information.....	51
Table 3.5: List of the selected speakers from the <i>Corpus of Kwansai Spoken Japanese</i> (Heffernan 2012) .....	53
Table 3.6: Comparative composition of heritage Japanese speakers and homeland speakers in the present study .....	54

### CHAPTER 4 VARIABLE REALIZATION OF SUBJECT PRONOUNS

Table 4.1: Summary of null subject languages .....	62
Table 4.2: Summary of rates of null subjects in various sources (spoken and written) .....	69
Table 4.3: Hypothesis for the effect of grammatical person on variant choice .....	83
Table 4.4: Hypothesis for the effect of subject continuity on variant choice .....	86
Table 4.5: Hypothesis for the effect of previous realization on variant choice .....	87
Table 4.6: Hypothesis for the effect of clause type on variant choice .....	87
Table 4.7: Hypothesis for the effect of verbal transitivity on variant choice .....	89
Table 4.8: Multivariate analysis of the contribution of social factors selected as significant to the probability of null subject pronouns .....	92

Table 4.9: Multivariate analysis of the contribution of linguistic factors selected as significant to the probability of null subject pronouns .....	94
--	----

Table 4.10: Overall rates of grammatical person in heritage Japanese and in the homeland benchmark variety .....	95
--	----

## CHAPTER 5 VARIABLE CASE MARKING ON NOUNS IN SUBJECT AND DIRECT OBJECT POSITION

Table 5.1: Hypothesis for the effect of verbal adjacency on variant choice .....	124
--	-----

Table 5.2: Hypothesis for the effect of presence of focus particle on variant choice .....	125
--	-----

Table 5.3: Hypothesis for the effect of presence of sentence-final particle on variant choice .....	128
---	-----

Table 5.4: Hypothesis for the effect of word order on variant choice .....	129
--	-----

Table 5.5: Hypothesis for the effect of clause type on variant choice .....	131
---	-----

Table 5.6: Overall rates of nouns for variable case marking in subject position .....	132
---	-----

Table 5.7: Overall rates of nouns for variable case marking in direct object position .....	133
---	-----

Table 5.8: Multivariate analysis of the contribution of social factors selected as significant to the probability of null marking on subject nouns .....	137
--	-----

Table 5.9: Multivariate analysis of the contribution of linguistic factors selected as significant to the probability of null-marking on subject nouns .....	139
--	-----

Table 5.10: Multivariate analysis of the contribution of linguistic factors selected as significant to the probability of null marking on English-origin subject nouns according to length of stay in Canada for heritage Japanese speakers .....	142
---	-----

Table 5.11: Multivariate analysis of the contribution of linguistic factors selected as significant to the probability of null marking on Japanese subject nouns according to age for homeland Japanese speakers .....	143
--	-----

Table 5.12: Multivariate analysis of the contribution of social factors selected as significant to the probability of null marking on direct object nouns .....	145
---	-----

Table 5.13: Multivariate analysis of the contribution of linguistic factors selected as significant to the probability of null marking on direct object nouns .....	147
---	-----

Table 5.14: Multivariate analysis of the contribution of linguistic factors selected as significant to the probability of null-marking on Japanese direct object nouns according to age for homeland Japanese speakers..... 149

Table 5.15: Summary of findings on the variable case marking between varieties of Japanese ..... 150

## LIST OF FIGURES

### CHAPTER 1 INTRODUCTION

Figure 1.1: Comparison for variable realization of subject pronouns ..... 15

Figure 1.2: Comparison for variable case marking..... 17

### CHAPTER 2 LITERATURE REVIEW

Figure 2.1: Hierarchy of borrowability according to the intensity of language contact ..... 23

### CHAPTER 4 VARIABLE REALIZATION OF SUBJECT PRONOUNS

Figure 4.1: Overall rates of variable realization of subject pronouns..... 90

### CHAPTER 5 VARIABLE CASE MARKING ON NOUNS IN SUBJECT AND DIRECT OBJECT POSITION

Figure 5.1: Overall rates of null making on nouns in subject position and direct object position .. 134

## LIST OF ABBREVIATIONS

Glosses for the Japanese examples cited in this thesis may contain the following abbreviations.

ACC	Accusative
ASP	Aspect
AUX	Auxiliary
CAUS	Causative
COMP	Complementizer
COND	Conditional
COP	Copular
CT	Continuative
DAT	Dative
DECL	Declare
DO	Direct Object
GEN	Genitive
HON	Honorific
IMP	Imperative
INTJ	Interjectional
LNK	Linker
LOC	Locative
NL	Nominalizer
NOM	Nominative
NEG	Negation
PASS	Passive
PAST	Past
PERF	Perfective
PL	Plural
PRES	Present (Non-past)
PROG	Progressive
Q	Question Particle
REP	Repetitive
SE	Sentence Extender
SFP	Sentence-final Particle
STAT	Stative
SUB	Subject
TAG	Tag
TOP	Topic
VOL	Volitional

# CHAPTER 1

## INTRODUCTION

### 1.1 Background

Speakers in a language contact situation tend to use two (or more) languages on a regular basis, sometimes simultaneously within the same discourse. There is a vast literature discussing the outcomes of language contact from various linguistic perspectives such as generative grammar (e.g. Corrigan 2010; MacSwan 1999, 2005, van Coetsem 2000), psycholinguistics (e.g. Bullock and Gerfen 2004; Bullock and Toribio 2004; Polinsky 1995, 1997, 2006; Polinsky and Kagan 2007), historical linguistics (e.g. Baldi 1990; Hoenigswald 1960; Thomason and Kaufman 1988), sociolinguistics (e.g. Haugen 1950; Poplack and Levey 2010; Silva-Corvalán 1982, 1994; Weinreich 1953; Weinreich et al. 1968). Outcomes of language contact cover a wide range of disparate phenomena such as borrowing, codeswitching, convergence, transference, language attrition, and so on.

Some studies have investigated the role of social factors implicated in language mixing (Bentahila and Davies 1991, 1995; Gumperz 1977; Thomason 2001; Thomason and Kaufman 1988), while others have explored internal structural constraints conditioning language-mixing behaviour (e.g. Muysken 1995; Poplack 1980; Sankoff and Poplack 1981). In particular, the latter body of research focuses on certain linguistic phenomena and attempts to elucidate the underlying grammar that gives rise to surface patterns of variation in multilingual discourse.

Among the most contentious outcomes of language contact are those relating to code-switching and borrowing. Rather than categorizing these phenomena as distinct, some scholars see them as manifestations of related phenomena that can be conceptualized in terms of a continuum of language-mixing strategies (e.g. Bentahila and Davies 1991; Boyd 1993; Clyne 2003; Myers-Scotton, 2002; Thomason 2001; Treffers-Dallers 2005; Winford 2009). By contrast, other empirical studies attempt to demonstrate that borrowing and codeswitching are two clearly different mechanisms (e.g. Poplack and Meechan, 1998). Further details about the distinction between borrowing and codeswitching can be found in Chapter 2.

First language (L1) attrition has been widely investigated and discussed as an important issue resulting from language contact (e.g. Schmid 2002; Schmid et al. 2004; Seliger and Vago 1991). It is generally claimed that attrition can be attributed to the interlanguage effects of L2 on the L1 linguistic system (see Altenberg 1991; Dorian 1982; Lambert 1989; Pavlenko 2002; Romaine 1989; Seliger and Vago 1991, among others). According to Seliger and Vago (1991), L1 attrition is described as both *externally* induced linguistic change (e.g. transfer, interference, convergence) and *internally* induced linguistic change (e.g. generalization, simplification). When L1 input is limited in the L2 environment, followed by an increase in L2 fluency, L1 linguistic properties are superseded by those of L2. Therefore, if unexpected linguistic patterns are found in L1 discourse in contact situations, it tends to be concluded that those speakers' L1 is attriting due to language contact.

However, linguistic variability, which is greatly emphasized in this dissertation, is often uncritically attributed to change and attrition without considering the possibility that the observed variability may simply represent inherent variation in the language. Thus, it is necessary to investigate the monolingual homeland variety in order to see the underlying

grammar in non-contact situations. Comparison of a heritage language variety in contact scenarios with its monolingual homeland variety will enable us to identify whether the variability observed in the heritage variety is indicative of attrition due to language contact or whether it is the inherent variability seen in the monolingual variety.

Another major manifestation of language contact that has been a persistent source of controversy for more than a century relates to contact-induced structural change, or the achievement of grammatical similarity (convergence) between two or more languages as the result of sustained and intense contact. This contentious topic lies at the heart of the investigation pursued in the current dissertation. Much of the controversy centers on the issues of whether grammatical change has occurred in a contact setting, and, if so, whether it is the product of external causation or internal evolution (or some combination of both) (e.g. Thomason 2001; Poplack and Levey 2010; Poplack et al. 2012a).

Although claims of contact-induced change exist in the literature, closer examination of these claims reveals that, more often than not, they are based on superficial inspection of surface features reported to be shared by languages in contact. As Poplack and Levey (2010), Poplack et al. (2012b), and Travis and Torres Cacoullos (2015) emphasize, it is important to look beyond superficial traits shared by two or more languages in order to scientifically ratify claims of contact-induced change. This is because shared surface features, and even similar rates of occurrence of these features in contact varieties, may mask underlying structural divergence. It is for this reason that variationist sociolinguists attach fundamental importance to systematic examination of the social and linguistic conditioning underlying surface patterns of variation in contact scenarios. Close inspection of this underlying structure, or “grammar,” across varieties is essential in adducing evidence in favour of (or against) contact-induced change. Furthermore, it

is not the underlying variable structure of experimentally elicited data that constitutes the major object of study in variationist research, but rather data obtained from everyday spontaneous discourse recorded from speakers belonging to well-defined social groups. It is within this conceptual and methodological framework that the present study is situated.

## 1.2 Overview of the Present Study

The present study investigates the linguistic outcomes of Japanese-English language contact in Canada. Adopting a sociolinguistic variationist framework (e.g. Labov 1966, 1969; Sankoff and Labov 1985; Weinreich et al. 1968), I aim to assess whether or not Japanese spoken in Canada, referred to here as *heritage Japanese*, exhibits structural change due to prolonged contact with English. A crucial point of departure for assessing potential contact effects of English on heritage Japanese involves the identification and characterization of *conflict sites*, defined here as loci containing qualitatively or quantitatively divergent features in two languages in contact (Poplack and Meechan 1998:136). As explained in more detail in §1.5 and Chapters 4 and 5, the *conflict sites* that are targeted in this dissertation are located in three grammatical sub-domains: variable *realization of subject pronouns*, variable *case marking of subject nouns*, and variable *case marking of direct object nouns*, as illustrated in the following examples:

### (1.1) Realization of Subject Pronouns

#### a. Overt Subject Pronouns

*Atashi* wa i-e-na-i.  
 I TOP say-POT-NEG-PRES  
 ‘I cannot say [it].’

(CJECC/O102/180<sup>1</sup>)

---

<sup>1</sup> The number in parenthesis indicates a speaker’s identifier and line number or time counter where the example occurs in the corpus used in the present study.

## b. Null Subject Pronouns

$\emptyset$  *sonna koto mo shir-ana-katta-desu ne.*

SUB such thing even know-NEG-PAST-COP SFP

‘[I] did not even know such a thing.’

(CJECC/O106/25)

## (1.2) Case Marking of Subject Nouns and Direct Object nouns

a. *Horede, guntai ga bochibochi ne, toukaku o*

then military.force NOM gradually INT pre-eminence ACC

*arawashi-te-ki-ta.*

show-CT-COME-PAST

‘Then, the military force has shot up to preeminence gradually [in Japan].’

(CJECC/017a/203)

b. *Yoku sono hito  $\emptyset$  ne, chichana inu  $\emptyset$  tsure-te sanpo-shi-te-ta*

often that person NOM INT small dog ACC take-CT walk-do-CT-PAST

‘That person often took a walk with a small dog.’

(CJECC/010/281)

As seen in (1.1), the subject pronoun can be overtly realized or omitted in a sentence as Japanese is a pro-drop language. Furthermore, in Japanese, functional relations are indicated by postpositional particles such as the nominative case marker *ga* and the accusative case marker *o* as demonstrated in (1.2a). However, those case markers can be omitted, as in (1.2b).

Hagen and de Bot (1990:37) mention a difficulty in selecting which linguistic variables should be examined for language attrition. It has been generally agreed that some linguistic properties are more vulnerable to language change than others in language contact situations (e.g. Matras 1998, 2009). Bullock and Toribio (2004:92) state those grammatical properties which are especially vulnerable to convergence are precisely the ones that lie at the syntax and pragmatic/semantics interface. In addition, according to the Interface Hypothesis (e.g. Sorace 2005, 2011), linguistic domains situated at the interface between linguistic system (e.g. syntax)

and other cognitive domains (e.g. discourse or pragmatics) are more vulnerable to language attrition. The variable realization of subject pronoun and variable case marking are located within the domains involving the interface between (morpho-)syntax and discourse information (e.g. Fry 2003; Hinds 1983; Masunaga 1988; Nariyama 2000, 2003; Ono et al. 2000). Therefore, the three variables to be examined are ideal sites to investigate whether or not heritage Japanese is influenced by English. It may be the case that, due to contact with English, heritage Japanese spoken in Canada can undergo language change in the areas of grammar in question or that heritage Japanese speakers show possible attrition.

This study is based on naturalistic speech data collected via sociolinguistics interviews (Labov 1984) from Japanese-English bilingual speakers in Canada. The *Corpus of Japanese-English Contact in Canada* (CJECC) contains over 55 hours of digitally-recorded naturalistic Japanese-English bilingual conversation from 18 speakers in Toronto and 29 speakers in Ottawa. From the corpus, 16 speakers were selected as *heritage Japanese speakers* and used in this dissertation based on length of stay in Canada (i.e. speakers who have lived in Canada over 14 years) and sex. The majority of studies on heritage language, which have been conducted in the United States, define heritage languages as home languages (L1) spoken by early bilinguals in a restricted environment (Polinsky 2011:1). As proposed by Nagy (2015:310), on the other hand, the term *heritage language* in a Canadian context can be defined as a primary language (L1) that is neither of the official languages of Canada, English or French, nor an indigenous language. Following this definition in a Canadian context, all heritage Japanese speakers used in this dissertation were native speakers of Japanese. They were born and grew up in Japan, acquiring the foundation of Japanese grammar within the sensitive period, before they arrived in Canada. Further details of this corpus and the definition of heritage language are discussed in Chapter 3.

### 1.3 Contribution of Japanese-English Data to the Study of Contact-Induced Change

Japanese-English bilingual data have the potential to illuminate key issues in contact linguistics from at least two perspectives.

First, Japanese and English are typologically dissimilar languages. Previous studies suggest that areas of structural similarity between languages are important 'convening points' for structural convergence. For example, Thomason (2001:71) states that contact-induced change may "result in major typological changes in the borrowing language." She emphasizes the importance of typological distance between languages in contact and claims that structural borrowing is facilitated by the existence of typologically congruent structures rather than divergent ones<sup>2</sup>. Johanson (2002) introduces the notion of "attractiveness," which states that some structures are more susceptible to copying (i.e. borrowing) or to resisting influence according to the relative degree of structural affiliation between languages in contact. Bullock and Toribio (2004:91) consider convergence as "the enhancement of inherent structural similarities found between two linguistic systems," implying that convergence is less likely to take place between typological dissimilar languages. Winford (2003:98) also mentions that "the greater the congruence between syntactic structures in two language in contact, the greater the likelihood that one will replace the other." Summarizing, there is broad consensus that the more typologically similar two languages are, the easier it is for one language to borrow from or influence another. Conversely, typological incongruence is believed to be an impediment to the transference of other-language material in contact situations, although, according to Thomason (2001), this barrier is not an absolute one if contact is intense enough and sufficiently sustained.

---

<sup>2</sup> Thomason (2001) claims that there are no purely structural constraints on what can be transferred from one language to another, and that, given a propitious set of social conditions, anything may be borrowed.

These claims lead us to infer that in the case of Japanese and English, which are typologically dissimilar, opportunities for structural convergence should be reduced.

To verify this inference, it is instructive to turn initially to several studies of language contact involving Japanese and English. There are many studies on language contact between Japanese and English. Their focus is primarily on how bilingual speakers mix two languages in speech (e.g. Azuma 1987, 1993, 1997; Fotos 1995; Namba 2008; Nishimura 1985, 1986, 1995, 1997) as well as the influence of English on the Japanese wordstock (e.g. the adaptation of English loanwords, the use of English for lyrics in Japanese popular music and titles of movies or TV shows, and so on; see e.g. Loveday 1996; Nyman 2012; Stanlaw 2004). However, none of the previous studies on Japanese-English contact has investigated contact-induced structural change in Japanese using a systematic, empirically accountable approach. This is a challenge that the present study directly engages with. An overarching objective of the present dissertation is to elucidate whether or not Japanese in contact with English (i.e. heritage Japanese) exhibits evidence of contact-induced structural change.

#### 1.4 Typological Differences between Japanese and English

As mentioned in §1.2, variable realization of subject pronouns and variable case marking of subject and direct object nouns are examined as conflict sites between Japanese and English. Since Japanese diverges from English in a number of structural aspects, this section briefly presents some key structural features of Japanese. The first characteristic of the language involves canonical SOV word order whereas English is an SVO language. Since Japanese is a verb-final language, an object is followed by a verb. Observe the following example (1.3) of a Japanese sentence and its English equivalent.

- (1.3) *Chisako ga hon o yom-u*  
 (name) NOM book ACC read-PRES  
 ‘Chisako reads a book.’

Likewise, a nominal complement is followed by a copula in Japanese as seen in (1.4) whereas it precedes the copula *be* in English.

- (1.4) *Kore ga chizu da.*  
 this NOM map COP.  
 ‘This is a map.’

Although SOV is the basic word order, Japanese exhibits free word order before a sentence-final verb, a phenomenon referred to as “scrambling” (c.f. Saito 1983; Tsujimura 1996). The following sentences in (1.5a-e) show the flexibility of constituent order with the exception of the verb. Scrambling does not change the meaning of the sentence.

- (1.5) a. *Mary ga senshū kuruma de Kyoto ni i-tta.*  
           NOM last.week car by (city name) to GO-PAST
- b. *Senshū Mary ga kuruma de Kyoto ni i-tta.*  
     last.week           NOM car by (city name) to GO-PAST
- c. *Mary ga Kyoto ni senshū kuruma de i-tta.*  
       NOM (city name) to last.week car by GO-PAST
- d. *Kuruma de Mary ga senshū Kyoto ni i-tta.*  
     car by           NOM last.week (city name) to GO-PAST
- e. *Senshū Kyoto ni kuruma de Mary ga i-tta.*  
     last.week (city name) to car by           NOM GO-PAST  
     ‘Mary went to Kyoto by car last week.’

In Japanese, functional relations are expressed by postpositional particles (Martin 1975:38; Shibamoto 1985:124; Shibatani 1990; Tsujimura 1996:134-136, and so on). English does not have such features, and functional relations are identified by word order in English. Particles are

used for indicating the case relation of nominals: nominals in an argument position are marked by the relevant particle for case marking, such as nominative, accusative, etc. Observe the following sentence in (1.6). The subject *ooya-san* ‘landlord’ is marked by the nominative case particle *ga* while the direct object *inu* ‘dog’ is marked by the accusative case particle *o*. Although the presence of a particle is indicated in normative grammar, the particle may be variably omitted in actual speech.

- (1.6) *Ooya-san*            *ga*    *sugoku*   *ii*    *hito*    *de*    *saikin*   *inu*   *o*  
 landlord-Mr./Mrs.    NOM    very    good    person    COP.and    recently    dog    ACC  
*kai-hajime-ta*        *no*.  
 have-start-PAST    SFP  
 ‘My landlord is a very good person, and he got a dog recently.’  
(CJECC/003/22:39)

Some particles are used as postpositions that resemble prepositions preceding nouns in English (Johnson 2008:33; Tsujimura 1996:133-134). While case marking particles are used to refer to the grammatical function of the nouns as described earlier, the particles for postpositions have a meaning: for example, *ni* ‘to, in, at’, *e* ‘to’, *de* (‘at’), *kara* ‘from’, *made* ‘until’ and so on. Since the verb comes in sentence-final position in Japanese, a postpositional phrase appears before the verb, as shown in the following examples.

- (1.7) a. *Takashi*   *ga*    *Tokyo*        *kara*    *ki-ta*.  
 (name)    NOM    (city name)    from    come-PAST  
 ‘Takashi came from Tokyo.’
- b. *Shoko*   *ga*    *kissaten*    *de*    *hon*    *o*    *yom-u*.  
 (name)    NOM    coffee.shop    at    book    ACC    read-PRES  
 ‘Shoko reads a book at a coffee shop.’
- c. *Shoko*   *ga*    *kuruma*    *de*    *gakkou*    *ni*    *ik-u*.  
                   NOM    car            by    school    to    go-PRES  
 ‘Shoko goes to school by car.’

Some particles are postposed by a clause or a sentence, and they function as conjunctives

(c.f. Shibatani 1990):

- (1.8) a. *onaka ga sui-ta kara hiru-gohan o tabe-masu.*  
 belly NOM become.empty-PAST because lunch-meal ACC eat -DECL(PRES)  
 ‘**Because** I became hungry, I (will) eat lunch.’
- b. *onaka ga sui-ta ga tabemono ga nai.*  
 belly NOM become.empty-PAST but food NOM exist.NEG  
 ‘**Although** I became hungry, there is no food.’

Shibatani (1990) presents several more Japanese features that differ from English. For word order, the relative clause precedes the modified noun, as in (1.9), whereas it follows the modified noun in English. An interrogative sentence is made by attaching a particle *ka* to a verb, as seen in (1.10), without changing the word order, as often happens in English. Furthermore, in Japanese the complementizer *to* is positioned before a main verb, indicating the end of a complement sentence (1.11), while in English complementizers mark the start of a complement clause.

- (1.9) *Taro ga mi-ta eiga*  
 (name) NOM watch-PAST film  
 ‘The film that Taro watched.’

- (1.10) *Bangohan o tabe-ta ka?*  
 dinner ACC eat-PAST Q  
 ‘Did you eat dinner?’

- (1.11) *Watashi wa Kono hon ga omoshiroi to omou.*  
 I TOP this book NOM interesting COMP think  
 ‘I think **that** this book is interesting.’

Finally, sentence-final particles and interjectional particles, which appear at the end of the utterance (Hasegawa 2014; Iwasaki 1993) are also important characteristics of Japanese. According to Maynard (2002:100), particles are classified into two groups: particles with grammatical functions, as we already have seen above, and particles that are employed pragmatically to negotiate, *inter alia*, interpersonal relationships. Sentence-final particles and interjectional particles fall into Maynard's second type. Postpositional particles occurring in a sentence-final position or verb-final position are called sentence-final particles. Speakers' attitude towards their assertion can be represented by sentence-final particles (Kuno 1978; Martin 1975:914; Maynard 2002; Shibatani 1990). For example, the final particle *yo* is used to call the attention of the listener to the speaker's assertion, as in (1.12a), while the particle *kashira* expresses the speaker's doubt or uncertainty about his/her assertion, as in (1.12b). Other final particles are used to emphasize the speaker's assertion (e.g. *zo*), ask for agreement or confirmation from the listener (e.g. *ne*), and so on. Interjectional particles occur freely within a clause. These particles may be used to soften or adjust the speaker's speech or obtain the attention of the listener. The meaning of a sentence does not change regardless of the presence or absence of interjectional particles. The use of sentence-final particles and interjectional particles is prevalent in spoken Japanese (see Hasegawa 2001; Iwasaki 1993; Maynard 1997). Examples (1.13a-b) contain interjectional particles *sa* and *ne*. Interjectional particles can occur after an adverb (e.g. *kyou*, as in 1.13a), after an adverbial phrase and after a noun followed by the accusative case particle *o*, as in (1.13b).

(1.12) a. *Ashita wa hare da yo.*  
 tomorrow TOP sunny.weather COP SFP (you know)  
 ‘Tomorrow will be sunny, you know.’

b. *Ashita wa hare kashira.*  
 tomorrow TOP sunny.weather SFP(I wonder)  
 ‘I wonder whether tomorrow will be sunny.’

(1.13) a. *Kyou sa, eiga o mi-you.*  
 today INTJ film ACC watch-let’s  
 ‘Let’s watch a film today.’

b. *Uchi no mae de sa, neko o ne, mi-ta yo.*  
 house GEN front in INTJ cat ACC INTJ see-PAST SFP  
 ‘In front of (my) house, I saw a cat.’

Table 1.1 below presents a summary of key typological differences between Japanese and English.

Table 1.1: Summary of dissimilarities between Japanese and English

Japanese	English
OV	VO
CV(copula)	V(copula)C
Scrambling of word order	Scrambling is not allowed
NP + Postpositional particle (case marking)	No case marking
NP + Postposition (particles)	Preposition + NP
Clause + conjunctives (particles)	Conjunctives + Clause
Relative Clause + NP	NP + Relative Clause
Complement + V	V + Complement
Sentence + <i>ka</i> (for interrogatives)	Word order (for interrogatives)
Presence of sentence-final particles and interjectional particles	No such particles

## 1.5 Goals

### 1.5.1 Research Questions

The main objective of the study is to elucidate whether or not Japanese spoken in Canada (Heritage Japanese, hereafter) is showing structural change due to contact with English. Speech patterns in heritage Japanese are explored systematically, drawing on a sociolinguistic variationist framework (e.g. Labov 1966; Sankoff and Labov 1985). My primary research questions are as follows:

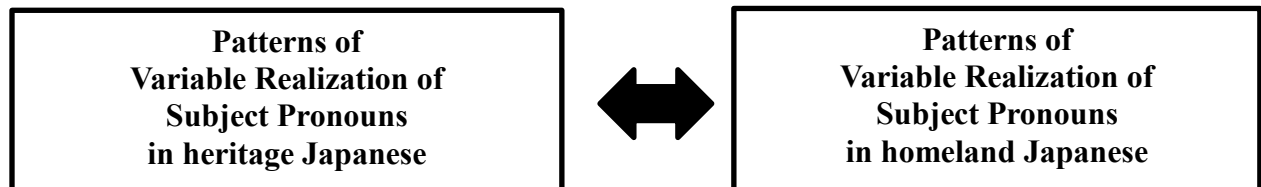
- What are the major sociolinguistic patterns of variation in the speech of bilingual speakers that are revealed by a quantitative analysis of targeted linguistic variables?
- What evidence is there that contact-induced structural change has occurred in heritage Japanese?

In order to answer these questions, I exemplify with three linguistic variables: variable *realization of subject pronouns* and variable *case marking on subject nouns* and variable *case marking on direct object nouns*. The patterns associated with these variables in heritage Japanese are systematically compared with a homeland baseline or control variety representative of contemporary spoken Japanese. In other words, the heritage variety is compared with a commensurate *vernacular* variety of Japanese. Justification for targeting the aforementioned linguistic variables is provided below.

### 1.5.2 Variable Realization of Subject Pronouns

In English, it is widely recognized that the subject of a sentence does not drop (i.e. non-pro-drop language), while in Japanese, the subject can often be null in conversation<sup>3</sup>. If heritage Japanese in Canada is influenced by English, it is hypothesized that there will be identifiable differences between null subjects in heritage Japanese and homeland Japanese as emerges from a systematic comparison of the underlying grammars. To test this hypothesis, patterns of realization of subject pronouns (i.e. overt or null) in heritage Japanese are compared with their counterparts in the homeland benchmark variety, as schematized in Figure 1.1. If heritage Japanese is behaving similarly to homeland Japanese, it is assumed that there is no contact-induced change in heritage Japanese with regard to the realization of subject pronouns. If heritage Japanese is behaving differently from homeland Japanese in terms of the social and linguistic conditioning of variability, the case for contact-induced change is strengthened, especially if targeted variable structures in the heritage variety can be shown to parallel in non-trivial ways linguistic patterns in the majority language, English.

Figure 1.1: Comparison for variable realization of subject pronouns



<sup>3</sup> Although English is recognized as a non-pro-drop language, null subjects in English have been reported in the literature (Harvie 1998, Haegeman 1997, Weir 2009, and among others). See Chapter 4 for details.

### 1.5.3 Variable Case Marking on Nouns in Subject and Direct Object Position

It has been widely reported that English nouns occur in Japanese-English bilingual discourse (Azuma 1997; Namba 2008; Nishimura 1997; Yoshizumi 2012). Using Japanese-English bilingual data in Canada, Yoshizumi (2012) found that lone English-origin nouns in subject and direct object position in otherwise Japanese discourse were adhering to the Japanese grammatical system with regard to variable case marking. Of particular concern in the present dissertation is whether case marking in heritage Japanese diverges from the system found in the monolingual Japanese spoken in Japan. As English employs no system of case marking analogous to that found in Japanese, it may be the case that heritage Japanese speakers' sustained contact with English will result in the erosion of their L1 case marking system. This is not an improbable scenario: the literature contains numerous studies of heritage varieties that are reported to experience structural attrition when a construction or linguistic system in the heritage language has no counterpart in the majority language (see Dorian 1981; Seliger and Vigo 1991; Polinsky 1995, 1997). Loss or simplification of case markers has been reported in other heritage languages in contact with English in the United States (e.g. Clyne 2003:124-130; Meyerstein 1969; Larmouth 1974; Preston 1986)<sup>4</sup>. Therefore, if English is influencing heritage Japanese in Canada, it is hypothesized that some identifiable differences between heritage Japanese and homeland Japanese will be observed by comparing their underlying grammars for variable case marking.

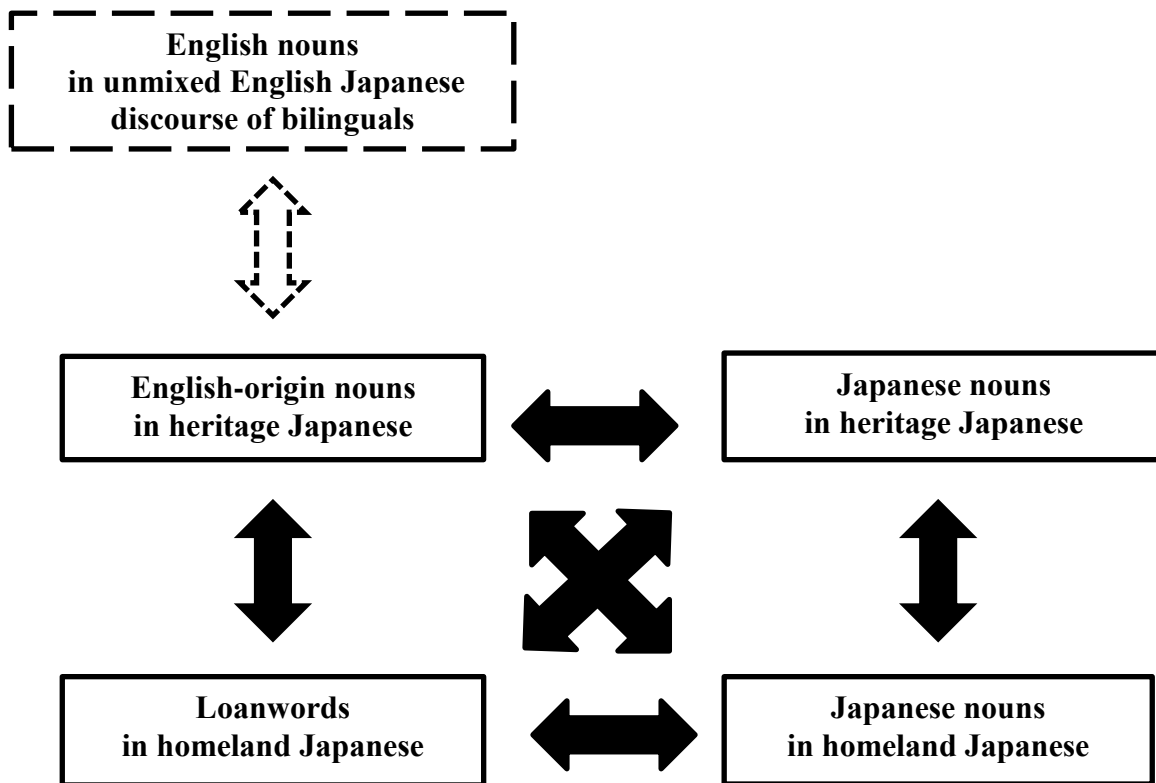
The present study investigates whether the underlying grammar for case marking of English-origin nouns and Japanese nouns in heritage Japanese discourse are the same as or

---

<sup>4</sup> However, Leisiö (2006) has reported a simplified case system in heritage Russian in contact with Finnish that has a more elaborate marking paradigm than Russian. Therefore, it may be that the simplified case system is not always caused by contact with a language without a case-marking system such as English.

different from the underlying grammar in homeland Japanese. Adopting the comparative method, English-origin nouns and Japanese nouns are compared with their counterparts in the monolingual homeland variety. The comparative axis of this component of the research is schematized in Figure 1.2.

Figure 1.2: Comparison for variable case marking



This approach, based on multiple comparison points, is a vital component in this study as none of previous research on English-Japanese bilingual data attempted consistent and systematic comparative analyses.

The comparative method adopted in the present study with regard to variable case marking also enables us to address the following questions:

- To what extent are English-origin nouns integrated into the otherwise Japanese discourse of heritage speakers?
- Is the case-marking pattern on English-origin nouns similar to or different from that on Japanese nouns in heritage Japanese?
- Are case-marking patterns on Japanese nouns and loanwords in homeland Japanese similar to or different from each other?
- Are English-origin nouns and Japanese nouns in heritage Japanese behaving similarly to or differently from Japanese nouns and loanwords in homeland Japanese?

Examination of these questions from a comparative perspective enables us to evaluate whether heritage Japanese is undergoing contact-induced change. If English-origin nouns and Japanese nouns in heritage Japanese are behaving differently in non-trivial ways from their homeland counterparts, then the inference that such differences are the product of contact-induced change is one that we pursue. Conversely, if we detect no substantial difference between the heritage and homeland comparison varieties, then the strength of any claim of contact-induced change is vitiated.

## 1.6 Theoretical Framework

The present study is situated within the theoretical framework of comparative variationist sociolinguistics (Labov 1963, 1966, 1969; Poplack and Tagliamonte 2001; Sankoff 1988; Sankoff and Labov 1985; Weinreich et al. 1968). Variationist theory holds the notion that “the

same linguistic *function* may at times be realized in different *forms*” (Poplack and Tagliamonte, 2001:7, their emphasis) and attempts to “discover usage patterns in the relative frequency of occurrence or co-occurrence of structures.” Under this approach, the language- or variety-specific variability is identified by investigating the systematic conditioning and distribution rate of variants, using naturalistic data.

The present study respects a number of key tenets and working principles of variationist sociolinguistics. First, quantitative analysis is an essential methodological tool as our research does not examine one particular variant or one-time occurrence of variants but recurrent variability. Patterns of variability are identified by examining rates and conditioning of variability. We also need to identify quantitatively what kinds of linguistic and extralinguistic factors influence speakers’ choice of one variant over the others. Quantitative analyses are an important prerequisite to elucidating patterns of inherent variability found in speech. Second, the comparative method (cf. Poplack and Tagliamonte 2001, 2013; Tagliamonte 2006) is adopted to explore the underlying grammatical relationship between heritage Japanese and homeland Japanese. Systematic comparison of the underlying grammars of these varieties enables us to identify similarities and dissimilarities between heritage and homeland Japanese.

## 1.7 Overview of the Dissertation

The ensuing sections of this dissertation are organized as follows:

Chapter 2 provides a discussion of contact-induced structural change and explores the distinction between borrowing and code-switching. It also offers a critical synopsis of previous research addressing Japanese-English language contact.

Chapter 3 present the theoretical framework and the methodology, including details pertaining to data collection, the sampling population, methods of extracting and coding tokens, data analysis, and so on.

In Chapter 4, I start with the literature review on null subjects in general and in contexts of language contact. I analyze the social and linguistic factors conditioning the distribution of variable realization of subject pronouns (overt or null). The underlying grammars on the realization of subject pronouns are then compared between heritage and homeland Japanese.

In Chapter 5, I review the literature on case marking in contexts of language contact, followed by detailed analyses of factors conditioning variable case marking in heritage Japanese. The underlying grammars of heritage Japanese is then compared with that of monolingual Japanese.

Finally, Chapter 6 presents general conclusions containing an overall discussion of the results obtained in Chapter 4 and 5, along with suggestions for further research.

## CHAPTER 2

### LITERATURE REVIEW

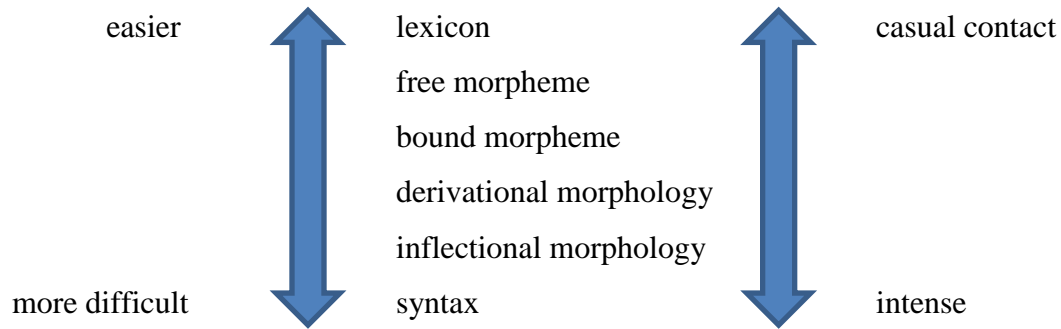
#### 2.1. Outcomes of Contact-Induced Change

Contact-induced grammatical change, or convergence, refers to the “achievement of greater structural similarity in a given aspect of the grammar of two or more languages” (Silva-Corvalán 2001:4-5). It is important to emphasize that this may involve a number of different processes. It may involve the borrowing and incorporation of other language material (e.g. the borrowing of English prepositions that are combined with French verbs to form new verb + preposition combinations in Acadian French, such as *parler about*, *aller on*, *faire up*; see King 2000). Or it may not involve the direct transfer of other language material (e.g. grammatical replication); instead, a new grammatical construction may arise in one language based or ‘calqued’ on a model construction in another language (see Heine and Kuteva 2010:86). Additionally, contact-induced structural change has been linked with the increase in the use of an existing minority variant in a language by virtue of contact with another language in which the equivalent structure is the majority variant. A good illustration of this kind of alleged contact-induced change can be found in US varieties of Spanish. In these varieties, it is claimed that overt subject pronominal expression, an existing minority option in Spanish, is increasing by virtue of intense and sustained contact with English, the majority language, in which overt subject pronominal usage is the norm (see Otheguy et al. 2007).

Speakers in a language contact situation tend to use two (or more) languages, and the regular use of two (or more) languages in contact scenarios has been equated with a range of

outcomes. Over the past several decades, these outcomes, particularly those associated with contact-induced structural change, have generated a great deal of controversy. In earlier studies, it was claimed that language change was an inevitable product of language contact (see e.g. Bynon 1977:240; Appel and Muysken 1987:154; Thomason and Kaufman 1988). Many historical linguists have postulated a hierarchy of structural constraints on what can be transferred from one language to another in contact scenarios (see Appel and Muysken 1987; Matras 2009:153-165; Moravcsik 1978; Weinreich 1953:35). According to these constraints, lexical items are the most labile linguistic component, and can be transferred easily from one language to another. By contrast, syntax and inflectional morphology are claimed to be most resistant to transference. It is also claimed that contact-induced change is predicted by social factors. It is generally considered that contact-induced change is more likely to occur when the degree of contact intensity is higher (Thomason 2001; Thomason and Kaufman 1988). Thomason and Kaufman (1998:74-76) proposed a borrowing scale, which displays a continuum from lexical borrowing to heavy structural borrowing based on *intensity of language contact*. According to this scale, lexical borrowing is the most easily observed in casual contact, whereas extreme structural change occurs in the most intensive contact settings. The constraints incorporating the intensity of language contact are summarized in Figure 2.1.

Figure 2.1: Hierarchy of borrowability according to the intensity of language contact



(Appel and Muysken 1987; Matras 2009; Moravcsik 1978; Thomason and Kaufman 1998; Weinreich 1953)

Other social factors that are considered as favourable to contact-induced change include length of language contact and level of bilingualism. A high level of bilingualism and a long period of contact between two languages have been reported to lead to (extensive) structural change (Thomason 2001; Thomason and Kaufman 1988; Winford 2003). These social factors shall be considered in the proposed study for assessing contact-induced change in Japanese spoken in Canada. Following the literature, it is hypothesized that speakers who have been in Canada for a longer period should be more likely to exhibit the effects of contact-induced structural change in their Japanese speech patterns than those bilingual speakers who have been residents in Canada for a shorter duration.

Turning to linguistic constraints on what may be transferred from one language to another, in spite of the conventional insistence on structural constraints, as depicted in Figure 2.1 above, Thomason (2001:71) maintains that there are no purely linguistic impediments to structural borrowing:

“Anything goes, including structural borrowing that results in major typological changes in the borrowing language. In phonology, loss or addition of entire phonetic and/or phonological categories in native words and of all kinds of morphophonemic rules. In syntax, sweeping changes in such features as word order, relative clauses, negation, coordination, subordination, comparison, and quantification. In morphology, typologically disruptive changes such as the replacement of flexional by agglutinative morphology or vice versa, the addition or loss of morphological categories that do not match in source and borrowing languages, and the wholesale loss or addition of agreement patterns. (Thomason 2001:71)

Thomason claims that contact-induced change may “result in major typological changes in the borrowing language” because “anything goes” for structural borrowing under the appropriate social circumstances. Thomason’s (2001) perspective clearly differs from the traditional position espoused by historical linguists, according to which syntax is the most resistant level of language to transfer.

However, recent empirical studies carried out within a variationist framework have argued that claims of dramatic structural change induced by language contact (as per Thomason 2001) are overemphasized (Poplack and Levey 2010; Poplack et al. 2012a, 2012b). A review of the relevant literature reveals that inferences of contact-induced structural change are widespread, but are often based on superficial parallels between languages in contact. According to Poplack and Levey (2010:398), these inferences invite criticism because: a) researchers often fail to verify whether a candidate linguistic feature for change in a contact variety is present in a relevant non-contact variety; and b) little attention is paid to whether or not the candidate feature for change in the non-contact variety behaves in non-trivial ways like its counterpart in the putative donor language.

Another methodological shortcoming, and one which the study in this dissertation addresses directly, is the nature of the variety that often serves as the benchmark with which

contact varieties are compared in order to identify and characterize change. All too often, the comparative baseline is some version of the standard variety of the contact language. A number of problems emerge when standard varieties are used to assess change in a contact variety. Firstly, standard languages are idealizations: they may never have corresponded to how a particular variety (contact or non-contact) was actually used by speakers in everyday speech. Secondly, standardization as a historical process tends to lead to the suppression of variability and the (artificial) promotion of invariability or fixation. When an inherently variable vernacular variety of a language is compared with an invariant or idealized standard, any deviations detected often tend to be interpreted as *changes*, and in bilingual settings, these alleged changes are, more often than not, attributed to contact. The problem with such comparisons (and the interpretations that ensue from them) is that they tend to overlook the fact that contact varieties of a language, just like their non-contact counterparts, tend to be variable too. Insufficient acknowledgement of the inherent variability of contact varieties can lead to the mistaken equation of language variation with language change. In other words, what might look like contact-induced change in a contact variety may also be explainable in terms of inherent variability. Thus, when looking at putative changes in heritage varieties, it is essential that these varieties are compared with a suitable non-contact *vernacular* control, and not a prescriptively ratified norm.

## 2.2 Distinguishing Borrowing and Codeswitching

Two of the primary (yet contentious) manifestations of language contact are *borrowing* and *codeswitching*. The distinction between borrowing and codeswitching has been hotly debated for a long time (e.g. Adalar and Tagliamonte 1998; Bentahila and Davies 1991; Boyd 1993; Boztepe

2003; Bullock 2009; Clyne 2003; Gardner-Chloros 2009; Johanson 1993; Meechan and Poplack 1995; Myers-Scotton 1993, 2002, 2006; Park 2006; Poplack and Meechan 1995, 1998; Poplack et al. 1988; Thomason 2003; Treffers-Daller 2005; Winford 2003, 2009). Some claim that a distinction between the two phenomena is not necessary (e.g. Bentahila and Davies, 1991; Boyd 1993; Boztepe 2003; Clyne 2003; Myers-Scotton 1993, 2002, 2006; Thomason 2003; Treffers-Daller 2005; Winford 2009), while others claim that borrowing should be unquestionably distinguished from codeswitching since they involve different mechanisms (Poplack and Meechan 1998). In fact, many studies have demonstrated that borrowed items and codeswitched items do not behave similarly in terms of the language system that they are following (Adalar and Tagliamonte 1998; Budzhak-Jones 1998; Budzhak-Jones and Poplack 1997; Eze 1997, 1998; Poplack 1993; Poplack and Dion 2012; Samar and Marjory 1998; Turpin 1998; Yoshizumi 2012).

The identification of the system membership of lone other-language items has often been a matter of debate due to methodological differences. However, a large number of empirical studies on data of various language pairs conducted within a variationist framework have demonstrated that lone items from a donor language pattern similarly to their counterparts in the recipient language. This has led to the classification of these items as borrowings (Poplack and Meechan 1998). Variationist theory holds the notion that “the same linguistic *function* may at times be realized in different *forms*” (Poplack and Tagliamonte 2001:7, their emphasis) and attempts to “discover usage patterns in the relative frequency of occurrence or co-occurrence of structures” (Poplack and Tagliamonte, 2001:7). Using this approach, language-specific variability is identified by investigating the systematic conditioning and distribution rate of variants, using naturalistic data. Even if non-standard behavior, as defined by the prescriptive norms of a language, is observed on lone other-language items in otherwise recipient-language

discourse, this is not enough evidence to conclude that the lone other-language items have failed to integrate into the grammar of the recipient language. Budzhak-Jones (1998) found that non-standard case-marking observed on lone English-origin nouns in otherwise Ukrainian discourse was reflective of the variability that exists in monolingual Ukrainian discourse. Given such inherent variability in the recipient language, recipient-language items should also be analyzed within the same corpus to see whether lone other-language items are following the grammatical system of the donor language or the recipient language.

It is hypothesized that a set of donor-language (Ld) items embedded in the recipient language (Lr) are borrowings if they have the same grammatical system as their Lr counterparts in the Lr discourse; on the other hand, a set of Ld items embedded in the Lr discourse are code-switches if they have the same grammatical systems as other Ld items and diverge from the grammar of their Lr counterparts in the Lr (unmixed) discourse (Poplack and Meechan 1998). This hypothesis is examined by identifying the recurrent patterns of Ld items in otherwise Lr discourse and comparing their counterparts in the unmixed discourse. This method enables us to determine the system membership of the Ld items in terms of relative degree of other language integration into Lr language. This variationist method is adopted in numerous studies which analyze bilingual data from typologically different language pairs.

For example, Poplack et al. (1989), which examines the validity of the equivalence constraint, demonstrated that borrowing is distinguished clearly from codeswitching in bilingual data of an SOV language (Finnish) and an SVO language (English)<sup>1</sup>. They found that lone English-origin nouns in otherwise Finnish discourse pattern like their Finnish counterparts,

---

<sup>1</sup> The equivalence constraint states that codeswitching is likely to “occur at points in discourse where juxtaposition of L<sub>1</sub> and L<sub>2</sub> elements does not violate a syntactic rule of either language” (Poplack 1980: 586).

appearing with obligatory case-marking or a function word. This finding indicates that these English-origin nouns are incorporated into the Finnish system and therefore are borrowings. In addition, regarding English-origin nouns that do not take Finnish obligatory case-marking, it was revealed that bare English-origin nouns were associated with functional flagging. This study demonstrates a clear distinction between borrowing and codeswitching.

Sankoff et al. (1990) also investigated the status of English-origin nouns found in Tamil-English data. Unlike English, Tamil is a SOV language and has postpositional case-marking. The authors examined the behaviour of English-origin nouns in object position in otherwise Tamil discourse by comparing their patterning of variability in case-marking (i.e. case-marked or unmarked) with native Tamil nouns in unmixed Tamil discourse. They found that variable case-marking patterns of English-origin noun complements governed by Tamil verbs were parallel to those of their native Tamil counterparts in unmixed Tamil discourse. As for variable accusative case-marking, both English-origin nouns in Tamil discourse and native Tamil nouns in unmixed Tamil tend to take case-marking in the obligatory marking context, whereas they are less likely to be case-marked in the optional marking context. These results indicate that even though English-origin nouns in otherwise Tamil discourse are not marked for the case, they are still behaving similarly to their native Tamil counterparts since their case-marking patterns are parallel to the variable case-marking patterns observed in Tamil. It is thus suggested that borrowed English nouns are integrated into the system of the recipient language (Tamil)<sup>2</sup>.

In sum, drawing on the comparative sociolinguistic method (Poplack and Meechan 1998), a key methodological step in assessing the status of “borrowed” items relies on multiple

---

<sup>2</sup> There are some more studies focusing on the variability of case marking on nouns as a diagnostic for determining to what extent lone English-origin nouns are incorporated into the system of the recipient language. See Eze (1997, 1998) for Igbo, Yoshizumi (2012) for Japanese, and so on.

comparisons between the behaviour of “borrowed” items in the recipient language and their corresponding behaviour in the donor language to assess which grammar is operative when a “borrowed” item is used in spontaneous discourse. Previous studies that have adhered to this methodology have revealed that lone English-origin nouns occurring in each recipient language are patterning like their recipient-language items, suggesting that they are borrowings that are integrated into the system of the recipient language. By conducting quantitative and systematic analyses of the patterning of lone other-language items, the system membership of these items can be determined unambiguously and consistently.

## 2.3 Language Contact between Japanese and English

### 2.3.1 Japanese-English Contact in Monolingual Settings

Language contact between Japanese and English, two typologically dissimilar languages, has been discussed from various perspectives. First, there is a great number of studies on the influence of English on Japanese spoken in monolingual contexts, focusing on the adaptation of English loanwords, Japanese people’s attitude toward English, and so on. Although the Japanese language has been historically influenced by various foreign languages such as Chinese, Korean, Portuguese, and Dutch, “English is the main donor-language in the contact processes taking place in contemporary Japan” (Loveday 1996:91). English lexical items are often concentrated in specific semantic fields: television, film and music industries, the brand name of commercial products, and so on (Harrmann 1989; Hoffer et al. 1983; Hoffer and Honna 1988; Loveday 1996; Nyman 2012; Stanlaw 1987, 2004). The influx of English-origin loanwords is reported to be relentless (Daulton 2008; Miller 1967; Shibatani 1990):

In fact, while Japanese phonology, syntax, pragmatics and discourse have remained relatively impervious to outside influence, a distant and exotic language -English- has extensively and fundamentally transformed the Japanese lexicon. (Daulton 2008:1)  
(my emphasis)

However, one of the problems with this type of claim is that it is made without reference to the principle of accountability (Labov 1969). In other words, lexical borrowings are not contextualized in relation to the native lexical items with which borrowings co-vary. For example, similar claims have been made in relation to the dramatic influence of English on the lexical resources of Canadian French; however, when these borrowings are contextualized in relation to native lexical items, accountable analyses of borrowing have shown that the reported dramatic influence of English on French has been exaggerated (Poplack et al. 1988).

It is widely acknowledged that use of English lexical items is highly prestigious in Japan (Backhaus 2006; Loveday 1996; Stanlaw 2004). English loanwords such as *chiketto* 'ticket' and *menyu* 'menu' are now widely used not only by bilingual speakers but also by monolingual Japanese speakers. Furthermore, Loveday's (1996) questionnaire-based survey shows that the degree of resistance toward contact with English is relatively low. Rather, acceptance and adoption of English innovations is favoured by certain demographic groups: those with higher educational background and higher occupation, and those aged between 18 and 29.

## 2.3.2 Japanese-English Contact in Bilingual Settings

In Japanese-English bilingual contexts, various English items have been observed.

- (2.1) *kyuu-sai kurai-no ko ni sura*, (E) “you look too young to be a nurse!”  
 nine-years.old about-GEN child by even

(J) *toka tte iwa-re-te.*

FP COMP say-PAS-CT

‘I was told by a girl about nine years old, “You look too young to be a nurse!” or something.’

(CJECC/014/42:33)

- (2.2) *Nihonjin ø nante bikkuri shi-te-ta.* (E) How come (J) *ne*  
 Japanese.people NOM FP surprise do-PERF-PAST ITJ

(E) you get English people to take you in there (J) *tte ne.*

COMP SFP

‘Japanese people were surprised like “How come you got English people to take you in there?”’

(CJECC/017a/48:02)

- (2.3) *Kanarazu teicyouna tegami ga ku-ru.* *Kanarazu* (E) a few line (J) *kai-te*,  
 always polite letter NOM come-PRES always write-CT

(E) Thank you very much (J) *dake ja-nai.*

only COP-NEG

‘A polite letter always comes. He always writes a few lines and it is not only “Thank you very much.”’

(CJECC/017b/1:00:21)

- (2.4) *Kodomo tte yuuno wa, hora, souyu imi de* (E) mean (J) *janai, sugoku.*  
 child COMP NM TOP INT that sense in TAG very

‘Children are very mean in that sense, aren’t they?’

(CJECC/012/5:02)

- (2.5) *Hito-sama no namae ø* (E) mistake (J) *shi-na-i.*

person-HON GEN name ACC do-NEG-PRES

‘He does not mistake someone’s name.’

(CJECC/018b/30:22)

In the examples in (2.1) and (2.2), the speaker introduces reported English speech. In (2.2), the Japanese interjectional particle *ne*, which is used to solicit agreement or confirmation from the listener, is inserted in the English quotation. English items can also occur at clause level, as in (2.3). An English nominal clause, ‘a few lines,’ occurs in an otherwise Japanese discourse. In (2.4), an English adjective, ‘mean,’ is followed by the Japanese copula *ja*, while in (2.5), an English verb, ‘mistake,’ is followed by a negative form of the Japanese pro-verb *shi-na-i* ‘do not (present)’.

Studies on Japanese-English contact in actual bilingual settings have been discussed for decades, predominantly focusing on codeswitching. Their primary focus is on investigating what kind of (other-language) items are *codeswitched*, with little attention paid to a distinction between codeswitching and borrowing (Azuma 1987, 1993, 1997; Namba 2008; Nishimura 1985, 1986, 1995, 1997). Nishimura (1985) and Azuma (1987) are the earliest studies that demonstrate patterns of intrasentential codeswitching in Japanese-English bilingual data. Their main argument centres on the validity of the equivalence constraint, which states that codeswitching tends to occur where two languages share the same constituent order (Poplack 1980; Sankoff and Poplack 1981). Both studies cast suspicion on the validity of the equivalence constraint and instead propose to account for Japanese-English mixed sentences by assigning a base language. According to Nishimura’s (1985, 1986) hypothesis “[w]hen switching takes place between constituents whose relative order is possible only in one language, that language is the language of the sentence” (1985:128); thus, the base language is determined by the constituent order.

However, the equivalence constraint has been wrongly interpreted by many previous studies on Japanese-English data (Azuma 1987; Nishimura 1985, 1995; Namba 2008 and others). The equivalence constraint does not restrict all codeswitching between two typologically

divergent languages such as Japanese and English. It stipulates that codeswitching is possible “at points in discourse where the juxtaposition of  $L_1$  and  $L_2$  elements does not violate a syntactic rule of either language” (Poplack 1980:586). Therefore, codeswitching between Japanese and English can occur where the constituent order is shared in discourse. This misinterpretation of the equivalence constraint by the researchers working on Japanese-English data is thought to be the result of disregarding a distinction between codeswitching and borrowing. If English items occur in otherwise Japanese discourse at points where the constituent order is not shared by Japanese and English, it is necessary to investigate further whether they are really codeswitched or something else (i.e. borrowing).

Nishimura (1997) and Namba (2008) demonstrate how other-language items are patterning by investigating to what extent the other-language items are integrated into the Japanese system or into the English system. However, they rely solely on prescriptive descriptions of Japanese to interpret the patterning of lone other-language items. In their account, for example, an English noun followed by a postpositional case-marking particle in otherwise Japanese discourse is considered to be integrated into the Japanese system. On the other hand, when an English noun is not followed by a case-marking particle, it is suggested that the noun in otherwise Japanese discourse retains the English system. The problem with this analytical procedure is that it invokes prescriptive norms as a benchmark. As discussed in §2.1, the prescriptively ratified benchmark cannot be used to draw systematic comparisons with a vernacular variety, as vernacular speech can diverge extensively from prescriptive norms, as confirmed by numerous sociolinguistic studies. The fact that the prescribed ‘obligatory’ case-marking particles may be variably omitted in monolingual vernacular Japanese serves as a cautionary reminder that comparisons with an idealized standard can be notoriously misleading. In fact, there are several

reports in the literature that prescribed obligatory case-marking particles can be dropped in colloquial speech (Martin 1975; Matsuda 1995; Murasugi and Sugisaki 2008; Ogawa 2008:787; Shibatani 1990; Takano 1998; Tsutsui 1984, and others). Nishimura (1997) and Namba (2008) investigate patterns observed in bilingual discourse only, but it would be impossible to determine whether or not the patterning of other-language items is the same as or different from the one observed in monolingual discourse without a scientific comparative analysis drawing on appropriate data.

Although many previous studies on Japanese-English contact have focused on the analysis of English loanwords and codeswitching between Japanese and English, little is known about whether or not there is any structural change in Japanese as a result of contact with English in everyday bilingual settings. This is a topic that the study described in this dissertation engages with paying careful attention to the inherent variability of everyday speech varieties.

# CHAPTER 3

## METHODOLOGY

### 3.1 Variationist Framework

The present study employs an empirical variationist sociolinguistic approach to assess contact-induced change in heritage Japanese in Canada. The fundamental premise of variation theory is that variation is inherent to human speech. This variability is not random or chaotic but exhibits structured heterogeneity. This ‘inherent variability,’ a hallmark of all human speech, is systematically constrained by a variety of internal (i.e. linguistic) and external (i.e. social) factors. The variability is typically conditioned not by a single contextual factor but by many contextual factors (i.e. the *principle of multiple causes* by Bailey in 2002). Variability typically exhibits form-function asymmetries, entailing that one form may encode multiple functions, and, likewise, that the same function may be encoded by multiple forms. The *linguistic variable* (Labov 1966) refers to a set of *variants* among which speakers alternate to express a given meaning or grammatical function. The choice of a particular variant over others belonging to the same system is influenced by linguistic factors as well as social factors. In order to investigate and interpret linguistic variability inherent in language, particular importance is attached to the following two methodological preconditions: the use of appropriate data and the *Principle of Accountability* (Labov 1972), as described in the following sections.

### 3.1.1 The Primacy of Vernacular Data

In the variationist framework, the use of appropriate data is a key requirement to investigate how speakers actually use language. Although non-sociolinguistic studies attempt to interpret language patterns using data collected from speakers' subjective grammatical judgements, by reading word/sentence lists, questionnaires, and so on, variationists foreground the importance of spontaneous speech. The optimal speech style for examining structured heterogeneity in everyday speech is the *vernacular*, which is defined as “the style in which the minimum attention is given to the monitoring of speech” (Labov 1972:208). Employing vernacular data is very important as the vernacular is the speech style which shows the most regular and systematic features in its linguistic structure (Labov 1972:112; Poplack 1993:252). The analysis of vernacular speech patterns offers the analyst a window on “real language in use” (Milroy 1992:66).

Variationists attempt to collect vernacular speech data, but the collection of such data is not without its problems. Chief among these are the effects which the researcher's/interviewer's presence may have on the speaker's production of casual, unmonitored speech. This gives rise to the *Observer's Paradox*, which is characterized by Labov (1972:92) in the following terms: “the aim of linguistic research in the community must be to find out how people talk when they are not being systematically observed; yet we can only obtain this data by systematic observation.”

To mitigate the effect of the Observer's Paradox and obtain a vernacular speech sample, variationists adopt a technique called the *sociolinguistic interview* (see Labov 1984). The sociolinguistic interview is designed “to approximate as closely as possible a casual conversation” (Schilling-Estes 2007:171). The interview strategy consists of a series of hierarchically

structured sets of questions within various conversational modules that are designed to elicit topics such as narratives of personal experience. In targeting such narratives, the expectation is that speakers become so involved in recounting an emotionally poignant experience (e.g. an accident or personal catastrophe) that they reduce their attention paid to speech, thus triggering a shift to vernacular usage (see Labov 1984).

### 3.1.2 The Principle of Accountability

The focus is on identifying and analysing recurrent patterns in discourse. A *pattern* can be defined as follows (Poplack and Meechan 1998:129):

- (3.1) A series of parallel occurrences (established according to structural and/or functional criteria) occurring at a non-negligible rate in a corpus of language use.

Patterns can only be identified systematically and quantitatively. As dictated by the *Principle of Accountability* (Labov 1969:737-738, fn. 20), the analysis must include “every case where the variable element occurs in the relevant environments as we have defined them,” which requires that “any variable form (a member of a set of alternative ways of saying the same thing) should be reported with the proportion of cases in which the form did occur in the relevant environment, compared to the total number of cases in which it might have occurred.” As Labov (1969) points out, unless this principle is followed, it is possible to prove any theoretical preconception by citing isolated or cherry-picked instances of a variable feature.

### 3.1.3. Quantitative Analysis

The variationist attaches a high premium to quantitative analysis. Quantitative techniques enable us to characterize the grammar underlying variant choice. The underlying grammar can be

revealed by the distribution and conditioning of competing variants in discourse (Poplack and Levey 2010:398). According to Sankoff (1988:84), we first perceive that speakers have a “choice between two or more specified sounds, words or structures during performance.” The choice process must occur repeatedly. Therefore, we focus on recurrent variability, not on one-time occurrence of variants.

The first step is to circumscribe the variable context in which the linguistic variable occurs. As dictated by the principle of accountability (Labov 1972), mentioned in the previous section, all variants (forms) are included in the relevant context, and each variant occurrence is recorded, additionally taking into account where a variant could have occurred but didn't. Any invariant context (i.e. contexts in which a variant occurs 100% or 0%) must necessarily be excluded from the analysis as these contexts do not admit any variation. Other contexts excluded from the analysis include speech disfluencies or anomalies, and false starts. From the well-circumscribed variable context, all relevant tokens of data are extracted. Since variant choice is considered to be conditioned by a number of linguistic and/or social factors, the extracted tokens are then coded for a number of different factor groups (or independent predictors) that are hypothesized to influence variant choice.

A distributional analysis is conducted to see how often the variants of a variable occur in a given body of data (Tagliamonte 2012:121), which enables the analyst to determine which factors, linguistic and extralinguistic, are associated with variant choice. In order to ascertain which factor groups, or independent predictors, are significant when all of them are considered simultaneously, the data are subjected to multivariate analysis, the choice analytical tool used by variationists (Tagliamonte 2006). Details of multivariate analysis are revealed through a series of probabilities and percentages and enable us to understand the grammar underlying variability

through “the three lines of evidence” (Poplack and Tagliamonte 2001; Tagliamonte 2006; Tagliamonte 2012). First, the *statistical significance* of a factor group contributing to variant choice is determined at the level of 0.05. Secondly, the *relative strength (magnitude)* of the factor groups can be assessed by the range value in each factor group. The range value is calculated by subtracting the lowest factor weight from the highest factor weight. It provides an indication of the strength of statistically significant factor groups and enables the relative magnitude of effect to be ranked in relation to one another in independent analyses. Thirdly, the direction of effect, or the order of factors from the largest to the smallest within independent predictors, which comprises favouring and disfavouring effects on the variant, makes up the *constraint hierarchy*. The constraint hierarchy can be understood as a window on a sector of the underlying grammar. We can make use of the information instantiated by the constraint hierarchy for comparative purposes. Where different speaker groups share structural similarities in the conditioning of variant usage, a relationship between the underlying grammars used by these groups can be inferred. On the other hand, where such correspondences are limited, or absent, evidence of an affinity between the underlying grammars is weakened.

### 3.2 The Comparative Method in Language Contact Studies

In the present study, assessment of contact-induced change is based on the comparative method (Poplack and Tagliamonte 2001; Tagliamonte 2012). This method has the capacity to illuminate the effects of possible contact-induced change in heritage Japanese by making use of three core constructs: (a) the conflict site, (b) a benchmark variety, and (c) the constraint hierarchy as the key comparative heuristic.

The notion of a *conflict site* is of particular utility in assessing contact-induced change. A conflict site exhibits structural and quantitative differences between two languages (Poplack and Meechan 1998:132). It is a key diagnostic for assessing whether the grammar of heritage Japanese is converging with that of English over time due to contact.

Secondly, assessing contact-induced change cannot be achieved without undertaking systematic comparative analyses. In the case of the present study, heritage Japanese spoken in Canada is compared with a monolingual contemporary variety of Japanese, spoken in Japan. Without a systematic comparison with the *homeland benchmark variety*, it is not possible to assess contact effects in heritage Japanese as we cannot investigate whether heritage Japanese speakers are exhibiting the same or different patterns in relation to speakers of the monolingual baseline.

In cross-varietal comparison, the *constraint hierarchy* is a key heuristic to assess contact-induced change. In much of the previous research, linguistic-specific patterns are identified by examining rates of variant occurrence as well as their sociolinguistic conditioning. However, frequency rates and the conditioning of variation do not have commensurate status: rates can vary for all kinds of extralinguistic reasons whereas the underlying grammar is expected to remain more constant, reflecting deeper and more stable constraints on variation (Poplack and Tagliamonte 2001:93-94). As emphasized in Poplack and Levey (2010:400) and Poplack and Tagliamonte (2001:92), differences in rates (or frequencies) of variant choice can only be used with caution to infer change. The conditioning of variability inferred from constraint hierarchies represents the underlying grammar and can be used to identify the relationship and provenance of forms.

By looking at whether or not the underlying grammar is shared between the languages in question, contact-induced change can be assessed “to the extent constraint hierarchies are language-specific” (Poplack and Levey 2010:400). In the multivariate analysis of datasets in the ensuing chapters, the constraint hierarchy is used as the principal heuristic for effecting systematic comparisons between varieties. If different varieties exhibit the same constraints on variability, as inferred from the constraint ranking, they are assumed to share the ‘same’ underlying grammar. To the extent that there are differences in the fine-grained patterning of the constraint ranking between different varieties, it can be inferred that such differences are indicative of the absence of a relationship between the grammars in question.

### 3.3 Japanese as a Heritage Language

Contact-induced change has been widely investigated in heritage languages, such as Spanish, Polish, Russian, Finnish, in contact with English in the United States. The general finding is that heritage language speakers exhibit different patterns from monolingual speakers of the same language. Recently, Laleko and Kawamura (2011) examined narratives elicited from heritage and monolingual speakers of Japanese, focusing on case-marking in Japanese, and found that heritage speakers underused the topic marker and overused the nominative case marker. Laleko and Polinsky (2013) additionally report that heritage Japanese speakers experience some difficulty in handling the use of the topic marker and the nominative case marker.

The present study also targets heritage Japanese, as spoken in Canada. However, the definition of ‘heritage language’ used in this study diverges in important ways from its use by other researchers such as Polinsky. Polinsky’s (2011:1, cited in Nagy 2015) definition is as follows:

Heritage languages are spoken by early bilinguals [...] whose L1 (home language) is severely restricted because of insufficient input. [...] they can understand the home language and may speak it to some degree but feel more at ease in the dominant language of their society.

(Polinsky 2011:1)

According to this definition, the heritage Japanese speakers described in Laleko and Polinsky (2013) are primarily engaged in *acquiring* the case marking system in Japanese. However, this is not the case for the heritage Japanese speakers in Canada examined in the present study. Nagy (2015:310) provides a definition of a heritage language associated with the Canadian context, drawing on Harrison (2000) and Cummins (2005). A heritage language in Canada is commonly construed as a primary language, or L1, that is neither of the official languages, English or French, nor an indigenous (First Nations) language. The heritage speakers who are targeted in the study described in the remaining chapters of this dissertation are all native Japanese speakers who were born in Japan, and who immigrated to Canada after adolescence. In other words, all these speakers had internalized Japanese as their L1 prior to arriving in Canada. Further details about the heritage speakers in this study are provided in the following chapter, which details the methodological framework underpinning the present investigation.

### 3.4 Data

#### 3.4.1 The Japanese-English Bilingual Community in Ontario (Toronto and Ottawa)

According to Statistics Canada (2012), as displayed in Table 3.1 below, the number of residents whose mother tongue is Japanese in the province of Ontario is approximately 12,215, which represents less than one third of the total number (39,985) in Canada. The province of Ontario contains the second largest number of Japanese mother-tongue speakers after British Columbia (15,840), where Japanese people first immigrated around the end of the 19<sup>th</sup> century.

Post-war immigration from Japan to Canada started to increase in the 1960s. Toronto, the provincial capital city of Ontario, is not only the city where a large number of Japanese-Canadians live, but also the one where a large number of Japanese people have immigrated for various reasons.

Table 3.1: Demographic profile of heritage Japanese

Japanese	Canada	Ontario	Toronto	Ottawa-Gatineau	British Columbia
mother tongue	39,985	12,215	6,230	1,030	15,840
spoken most at home	18,850	765	2,995	420	8,110
spoken regularly at home	19,050	5,765	2,780	595	6,870

(source: Statistics Canada, 2011 Census of Population)

A large number of Japanese companies established branch offices in Toronto, and many Japanese youth came for working-holiday programs and to study. There are various associations and organizations for the Japanese-Canadian community in Toronto, which has 6,230 mother-tongue-Japanese speakers according to Statistics Canada (2012), as indicated in Table 3.1 above. Some of these organizations exist to support Japanese people new to the community. The *Japanese Canadian Cultural Centre (JCCC)* is the largest centre for Japanese-Canadians, newcomer Japanese immigrants, and those of non-Japanese ancestry who are interested in Japan and its culture. The *JCCC* has played an important role in connecting people in the Japanese community by offering various cultural courses and events, organizing exhibitions of Japanese arts and film festivals, and so on.

Likewise, many Japanese-Canadians and recent Japanese immigrants reside in the National Capital Region, which consists of Ottawa, Gatineau (the province of Quebec), and their surrounding areas. Ottawa is the nation's capital and hosts Japanese diplomats and their families,

Japanese people who have married Canadians, as well as Japanese academic researchers. Statistics Canada (2012) reports that the total population of native Japanese speakers is 1,030, as indicated in Table 3.1 above. The Japanese community first consists of the *Ottawa Japanese Community Association (OJCA)* and the *Ottawa Japanese Cultural Centre (OJCC)*. They are sister organizations and promote Japanese culture in the area, similarly to the *JCCC*. The Japanese community is strongly connected to the embassy of Japan, which provides its own network of speakers of Japanese. Furthermore, there are quite a few Japanese researchers who immigrated to Ottawa to work at the National Research Institute (NRI). Those researchers started to immigrate to Ottawa from around the 1960s.

Furthermore, in both cities, Toronto and Ottawa, there are Japanese Saturday schools (called *hoshuko* or *hoshu-jugyoko*, translated as “a supplementary school” in English), which conduct classes such as Japanese, Mathematics, and so on, using the same textbooks as schools in Japan. These schools are supported by the Ministry of Education, Culture, Sports, Science and Technology associated with the government of Japan. Children of Japanese families have the opportunity to take classes in Japanese and enjoy Japanese cultural events similar to those offered by schools in Japan.

Established Japanese communities in both cities are ideal for studying Japanese-English bilingualism, and constitute an ideal venue for investigating the outcome of Japanese-English contact as well as, more specifically, for assessing contact-induced change in Japanese in a predominantly English-speaking environment.

### 3.4.2. Data and Speaker Sample

The speech patterns of Japanese spoken in Canada (hereafter referred to as heritage Japanese) will be systematically compared with those of a non-contact vernacular benchmark, monolingual Japanese spoken in Japan, termed *homeland Japanese*. The data were collected by employing the technique of the *sociolinguistic interview* (Labov 1984) in order to capture the patterns of inherent variability observed in heritage Japanese. Details of both data sets, which form the primary comparison points in this study, are given below.

The first data set used in the present study comprises the *Corpus of Japanese-English Contact in Canada (CJECC)*, which contains recorded spoken Japanese in Canada. The corpus consists of over 55 hours of digitally-recorded naturalistic Japanese-English bilingual conversation recorded from 18 speakers in Toronto and 29 speakers in Ottawa during the period of 2009-2013. As mentioned in 3.1.1, the type of data we investigate is the vernacular. Although all speech contains some style shifting, the data are generally representative of spontaneous usage. In the aggregate, the data are all relatively informal, as exemplified by the occurrence of code-switching, borrowing, the use of non-canonical word order, use of sentence-final particles, and so on.

Tables 3.2 and 3.3 present information about heritage Japanese speakers included in the corpus. Most speakers are post-war Japanese immigrants who came to Canada at some point during post-adolescence<sup>1</sup>. One-on-one sociolinguistic interviews were conducted with 40

---

<sup>1</sup> However, speaker T005, shown in Table 3.2, whose mother is Japanese and father is Australian, was born in Australia. After marrying a Canadian, she moved to Toronto. Speaker T017 was born in the United States; however, he moved back to Japan when he was little, after his parents passed away, and grew up in Japan. After working in the Japanese military, he immigrated to Vancouver, Canada, at the age of 18. Speakers T006, T015, and O113 are Canadian-Japanese who were born and grew up in Canada.

speakers, and an in-group interview was done with four married couples<sup>2</sup>. During the interview, I (the interviewer) typically used only one language, Japanese or English. The language used was selected depending on which language had been used in previous interactions with the interviewees. The speakers with whom I used English were the Japanese-Canadian residents belonging to the second generation (*'ni-sei'*) only<sup>3</sup>.

---

<sup>2</sup> One of the in-group interviews was done with speaker T002 and her Canadian husband who is a native English speaker. Her husband is not included in the *CJECC*.

<sup>3</sup> Note that these Japanese-Canadian speakers in the second generation do not adhere to the definition of heritage Japanese speakers in the present study.

Table 3.2: Information about heritage Japanese speakers in Toronto in the *CJECC*<sup>4</sup>

Speaker Number	City <sup>5</sup>	Interview Time (Total/h:m:s)	Hometown	Sex	Age	Years of residence	Education	Occupation	Language used in the interview
T001	TO	55:04	Tokyo	F	35	10	PS	Biologist	Japanese
T002	TO	1:23:45	Fukuoka	F	64	25	PS	Retired (Employee)	Japanese
T003	TO	1:18:59	Kumamoto	F	28	1.5	PS	Translator	Japanese
T004	TO	1:01:05	Chiba	M	70	35	PS	Retired (employee)	Japanese
T005	TO	19:53	Sydney, Australia	F	27	3	-	JCCC stuff	Japanese
T006	TO	1:14:05	Vancouver	M	76	76	S	Retired	English
T007	TO	40:47	-	F	30	3	PS	Student, JCCC stuff	Japanese
T008	TO	2:13 + 48:55	Gunma	F	31	3	-	Works at a coffee shop	Japanese
T009	TO	56:38	Tochigi	F	38	6	S	Nurse	Japanese
T010	TO	1:06:17	Chiba	F	45	16	-	Unemployed (former company employee)	Japanese
T011	TO	56:29	Miyazaki	F	57	33	S	Housewife, Helps her husband's company	Japanese
T012	TO	42:15	Niigata	M	65	39	PS	Self-employed	Japanese
T013	TO	31:58	Machida	F	38	15	PS	Health Canada	Japanese
T014	TO	1:03:02	Saitama	F	33	5	-	Nurse	Japanese
T015	TO	59:23	Vancouver	M	91	91	PS	Retired vice-president of a large commercial organization	English
T016	TO	1:10:35	-	F	35	1.5	-	Engineer	Japanese
T017	TO	3:09:56	U.S.	M	93	75	S	Retired (Self-employed)	Japanese
T018	TO	1:06:46	Iwate	M	54	23	PS	Self-employed	Japanese

<sup>4</sup> Specific information is withheld to protect the anonymity of the interviewee, in line with ethical requirements.<sup>5</sup> TO refers to Toronto. OT refers to Ottawa.

Table 3.3: Information about heritage Japanese speakers in Ottawa in the *CJECC*

Speaker Number	City	Interview Time (Total/h:m:s)	Hometown	Sex	Age	Years of residence	Education	Occupation	Language used in the interview
O101	OT	56:34	Hiroshima	F	70	41	PS	Retired (Government employee, Teacher)	Japanese
O102	OT	1:29:29	Yamanashi	F	69	39	PS	Retired (Restaurant owner)	Japanese
O103	OT	53:16	Tokyo	F	45	22	PS	Housewife, Volunteer	Japanese
O104	OT	1:41:37	Tokyo	F	62	29	PS	Housewife, Former employee at various companies	Japanese
O105	OT	59:30	Tokyo	M	41	15	PS	Company employee	Japanese
O106	OT	2:32:32	Tokyo	M	83	49	PS	Retired (Scientist at the NRC)	Japanese
O107	OT	1:25:09	Wakayama	M	65	41	-	Employee	Japanese
O108	OT	1:25:12		M	55	12	PS	Church minister	Japanese
O109	OT	1:25:12	Osaka	F	49	12	-	Housewife	Japanese
O110	OT	1:13:38	Kumamoto	M	55	27	PS	Employee	Japanese
O111	OT	2:05:46	Tokyo	M	82	51	PS	NRI researcher,	Japanese
O112	OT	1:31:09	Shiga	M	65	41	S	Automobile mechanic (Owner)	Japanese
O113	OT	56:36	B.C	M	92	92	-	Retired (Canadian Military)	Japanese
O114	OT	1:19:11	Hokkaido	F	72	44	-	-	Japanese
O115	OT	1:19:11	Tokyo / Osaka	M	72	46	PS	Federal government	Japanese
O116	OT	1:42:06	Osaka	M	66	42	S	Engineer (Owner)	Japanese
O117	OT	1:42:06	Osaka	F	66	42	-	-	Japanese
O118	OT	1:03:08	Kumamoto	F	44	14	PS	Housewife, Works at a day nursery	Japanese

Table 3.3 (continued from the previous page)

Speaker Number	City	Interview Time (Total/h:m:s)	Hometown	Sex	Age	Years of residence	Education	Occupation	Language used in the interview
O119	OT	54:31	Tokyo	F	40	16	PS	Canadian Food Inspection Agency	Japanese
O120	OT	1:16:55	Kumamoto	F	41	18	PS	Employee	Japanese
O121	OT	1:30:23 + 13:13	Hakodate	F	79	46	-	Tea instructor	Japanese
O122	OT	1:06:13	Kyoto	M	52	26	PS	Translator	Japanese
O123	OT	56:02	Kobe	F	49	20	PS	Florist, tour guide	Japanese
O124	OT	1:38:53	Shanghai /Tokyo	F	91	44	PS	Retired	Japanese
O125	OT	1:05:50	Saitama	M	57	36	S	Self-employ	Japanese
O126	OT	2:25:39	Saitama	F	45	-	-	Florist	Japanese
O127	OT	47:00	Tokyo	F	45	10	-	Housewife	Japanese
O128	OT	56:22	Nagoya	F	48	15	PS	Government employee	Japanese
O129	OT	58:48	Sendai	F	65	45	S	Hairdresser	Japanese

For the present study, 16 speakers were subsampled from the corpus, based on their length of stay in Canada and their sex. It is important to measure speakers' length of stay in Canada, because, as Thomason (2013:37) points out, "the longer the contact period and the greater the level of bilingualism, the more likely it is that structural features will be transferred along with lexical items." Since speakers' level of English ability was not evaluated at the time of the interviews, their length of stay in Canada is considered as a proxy of their length of contact period with English. The length of stay in Canada ranges from 14 years to 49 years (the median is 33.38 years). Regarding the years of stay in the host dominant language environment, De Bot

et al. (1991), which investigated the L1 maintenance or loss of Dutch immigrants in France, found no significant effects of the elapsed time since emigration among speakers. De Bot and Clyne (1994) also found no evidence of attrition over the 16 years, stating that the L1 attrition is not an inevitable outcome for immigrants.

Table 3.4 displays a breakdown of the speakers participating in the present study by average number of years of residence in Canada and provides more detailed information about the speakers, including demographic factors. Overall, the speaker sample is divided into two broad cohorts based on average years of residence in Canada: 20.375 years for those with the shortest residency, and 40.375 years for those with the longest, which doubles the former cohorts. As shown in Table 3.4, all the selected heritage Japanese speakers arrived in Canada in adulthood. They were born and grew up in Japan. They acquired the foundations of Japanese, their primary language, in Japan, within what is considered the sensitive period, and emigrated to Canada after this critical stage in language acquisition (i.e. De Bot et al. 1991:88). In addition, once in Canada, they have been exposed to English-speaking or Japanese-English bilingual settings on a daily basis due to work or family requirements. For example, they work or worked for the government, or an English-speaking company, or where advanced English proficiency is required. The speakers range in age from 41 to 55 for those with shorter residence in Canada and from mid-60s to 83 (except for one who is 57 years old) for those with longer residence in Canada.

Table 3.4: Composition of heritage Japanese speakers in Canada according to years of residence in Canada with more detailed information

	Speaker Number	Sex	Age	years of stay in Canada	Year of arrival	Age of arrival in Canada	Occupation
<b>14~27 years: Average <u>20.375</u> years</b>	O105	Male	41	15	1998	26	Company employee
	T018	Male	54	23	1986	31	Self-employed
	O122	Male	52	26	1987	26	Translator
	O110	Male	55	27	1986	28	Employee
	O118	Female	44	14	1999	30	Housewife, Works at a day nursery
	T010	Female	45	16	1993	30s	Unemployed (former company employee)
	O123	Female	49	20	1993	29	Florist, tour guide
	O103	Female	45	22	1991	23	Housewife, volunteer, (former company employee)
<b>33~49 years: Average <u>40.375</u> years</b>	T004	Male	70	35	1974	35	Retired (Company employee)
	T012	Male	65	39	1970	26	Self-employed
	O112	Male	65	41	1972	24	Automobile mechanic (Owner)
	O106	Male	83	49	1964	35	Retired (Scientist at the NRC)
	T011	Female	57	33	1976	23	Helps her husband's company
	O102	Female	69	39	1974	30	Retired (Restaurant owner)
	O101	Female	70	41	1972	30	Retired (Government employee, Teacher)
	O121	Female	79	46	1967	33	Tea instructor

In order to make systematic comparisons between heritage Japanese and monolingual vernacular Japanese, the present study draws on the *Corpus of Kansai Spoken Japanese* (Heffernan 2012) housed at the Kansai Gakuin University, Japan. The corpus consists of naturalistic conversation recorded from over 50 speakers, employing a sociolinguistic interview methodology (Labov 1972). This urban mainstream Japanese vernacular is used as non-contact control benchmark. For the present study, 16 speakers were selected based on speaker's age, sex, and their self-assessment of English ability. The sample of homeland Japanese speakers was constructed as shown in Table 3.5. The speakers' age range is from 31 to 80 years old. As selected speakers of heritage Japanese came to Canada at some point during post-adolescence and have stayed for more than 14 years, the youngest speakers are in their 30s. None of them is in their 20s. From the *Corpus of Kansai Spoken Japanese*, therefore, only speakers over 30 years old were selected. All of the 16 speakers report no English ability.

Table 3.5: List of the selected speakers from the *Corpus of Kwansai Spoken Japanese* (Heffernan 2012)

Speaker Code	Gender	Age	Occupation	Education	English Ability
026	Male	30-39	bus driver	high school	None
027	Female	30-39	Service	high school	None
035	Male	30-39	bus driver	high school	None
038	Male	30-39	bus driver	high school	None
039	Male	30-39	company employee	Junior college	None
024	Female	40-49	Housewife	Junior college	None
025	Female	40-49	Housewife	Junior college	None
047	Male	40-49	company employee	high school	None
030	Female	50-59	company employee	Junior college	None
037	Female	50-59	childcare worker	Junior college	None
017	Female	60-69	None	high school	None
013	Male	70-79	Retired (company employee)	junior high school	None
022	Male	70-79	Retired (company employee)	high school	None
023	Male	70-79	Retired (company employee)	high school	None
052	Female	70-79	/	junior high school	None
053	Female	70-79	Housewife	high school	None

Table 3.6 depicts a comparative summary of the heritage Japanese speakers and homeland speakers included in the present study.

Table 3.6: Comparative composition of heritage Japanese speakers and homeland speakers in the present study

Heritage Japanese speakers (Canada)				Homeland Japanese speakers (Japan)			
Length of stay in Canada	Male	Female	Total	Age	Male	Female	Total
Short (14~30 years)	<b>4</b>	<b>4</b>	8	Young (30~49 years)	<b>4</b>	<b>4</b>	8
Long (31~50 years)	<b>4</b>	<b>4</b>	8	Old (50~79 years)	<b>4</b>	<b>4</b>	8
<b>Total</b>	<b>8</b>	<b>8</b>	<b>16</b>	<b>Total</b>	<b>8</b>	<b>8</b>	<b>16</b>

### 3.5 Procedure

The present study adopts a comparative variationist framework, which has been described in the previous sections, to uncover the speech patterns in heritage Japanese in Canada and assess the effects of possible contact-induced change on heritage Japanese due to English influence. Within this framework, language-specific variability is identified by investigating the systematic conditioning and distribution of variants, drawing on naturalistic speech data.

Once the target variables have been extracted and coded, they will be analyzed using GOLDVARB X (Sankoff et al. 2012). By comparing and contrasting the patterning of variable case-marking and variable realization of subject pronouns taken from the variable rule analyses, similarities and dissimilarities between heritage Japanese and homeland Japanese can be revealed systematically and objectively. As stated in the section 3.2, constraint hierarchies reflect the underlying grammatical conditioning of linguistic variability. The comparison of constraint hierarchies will point to whether or not the same underlying grammar is operative in the two

varieties of Japanese. Findings obtained from the comparison of constraint hierarchies between heritage Japanese and homeland Japanese are used as the key metric for assessing contact-induced change.

## CHAPTER 4

### VARIABLE REALIZATION OF SUBJECT PRONOUNS

#### 4.1 Introduction

The first linguistic variable which is targeted with a view to investigating a possible contact-induced change is the variable *realization of subject pronouns* in Japanese. Japanese is well-known as a (radical) pro-drop language (e.g. Camacho 2013; Roberts and Holmberg 2010). The subject of a sentence can be overtly expressed or unexpressed as illustrated in the following examples:

(4.1) ***Watashi*** *Burajiru ni to-n-jaun-desu.*

I        Brazil    to    fly-LNK-AUX-COP(PRES).  
'I fly to Brazil.'

(CJECC/T004/240)

(4.2) ***Boku*** *Pat to isshoni kae-ru yaro.*

I        Pat    with together return-PRES    TAG  
'I go back [to Japan] with Pat, right?'

(CJECC/O112/155)

(4.3)  **$\emptyset$**  *hitori de ki-ta-n-desu.*

SUB    one.person by come-PAST-LNK-CPL  
'[I] came by myself.'

(CJECC/O106/136)

(4.4)  **$\emptyset$**  *dekai ne.*

SUB    tall    SFP  
'[He] is tall.'

(CJECC/O105/5)

- (4.5) *Sorede mou, Ø [Ø nidoto ii-taku-na-i na] to omo-u.*  
 then any.more SUB SUB twice say-DES-NEG-PRES SFP COMP think-PRES  
 ‘Then, [I] think that [I] don’t want to say (it) anymore.’  
 (CJECC/O102/178)

- (4.6) *Ø iroiro kurou-shi-ta hito-tachi da kara, gyakuni, sono,*  
 SUB various suffer-do-PAST person-PL COP because conversely INTJ  
*kare-ra-wa ishikitekini nihon-o sute-you to.*  
 hi-PL-TOP intentionally Japan-ACC abandon-VOL COMP  
 ‘Because [they] are the people who experienced all sorts of hardships, they tried to  
 abandon Japan intentionally.’  
 (CJECC/T012/132)

In (4.1-2), the subjects are expressed: *watashi* and *boku* are first person singular variants. In (4.3-4), the subjects are not expressed; however, from the ambient discourse context, the referent of the subject is interpreted as ‘I’ in (4.3) and ‘he’ in (4.4). In (4.5), the subject in the main and embedded clause is not expressed, but the referent ‘I’ is recoverable from the ambient discourse. In (4.6), the subject in the subordinate clause is not expressed while the subject in the following main clause is overtly expressed. By contrast, English is considered as non-pro-drop language, requiring that the subject pronoun be ordinarily overt. Therefore, variable realization of subject pronouns (i.e. overt or null) qualifies as a suitable conflict site (see Poplack and Meechan 1998:132) between Japanese and English. As such, any discernible increase in the use of overt subject pronominal realization in heritage Japanese, influenced by the prevalence of overt subject pronouns in English, and a concomitant decrease in the use of the null variant in heritage Japanese, might be indicative of contact-induced change.

The main objective in this chapter is to investigate the hypothesis of contact-induced change in relation to the variable realization of subject pronouns in heritage Japanese. By comparing the patterns of heritage Japanese with those of the homeland benchmark variety using

an empirically accountable quantitative approach, we will assess whether English is influencing the grammatical structure of heritage Japanese. It is hypothesized that if contact-induced change is present, then its effects should reside in a difference in: (a) rates of variant occurrence in heritage Japanese and homeland Japanese, as well as, crucially, (b) differences in the grammar underlying the observed variability.

The remainder of this chapter is organized as follows: the literature pertaining to null subject languages is reviewed in §4.2-§4.5 and then section 4.6 describes the methodological procedures adopted in the present study to circumscribe the envelope of variation and extract the target variable. Section 4.7 presents the hypothesis for the investigation. This is followed in § 4.8 by a discussion of the social and linguistic factors operationalized in the present study in order to investigate the possibility of contact-induced change. Section 4.9 presents the results of the investigation, and section 4.10 reviews the pertinence of the major findings with regard to the research objectives stated above.

## 4.2 Null Subject Languages

### 4.2.1 Null Subject Languages in the Earlier Studies

In this section, we present an overview of null subject languages in generative grammar. Null subject languages are those languages that allow the grammatical subject of a clause to be implicit, as in the following examples:

- (4.7) Spanish  
 $\emptyset$  *Salimos.*  
 left  
 ‘[We] left.’

(source: Camacho 2013:14)

- (4.8) Quechua  
 $\emptyset$  *Papa-ta*            *mikhu-n-mi*.  
 potato-ACC            eat-3SG-FOC/EVID  
 ‘[S/he] eats potatoes.’

(source: Sànchez 2004:24)

- (4.9) Chinese  
 $\emptyset$  *Kanjian ta le*.  
 see            him LE  
 ‘[He] saw him.’

(source: Huang 1984:533)

In the classical formulation of the null subject parameter, the following cluster, or set, of properties have been proposed (c.f. Chomsky 1981; Chomsky and Lasnik 1977; Perlmutter 1971; Taraldsen 1978): (a) the availability of null subject *pro*; (b) free inversion (postverbal subjects); (c) apparent violations of *that*-trace effects; (d) a rich verbal morphology.

Perlmutter (1971:100) is the initial study that introduced a typological distinction between null subject languages and non-null subject languages, proposing the following surface structure constraint:

- (4.10) Any sentence other than an imperative in which there is an S that does not contain a subject in surface structure is ungrammatical.

(Perlmutter 1971:100)

According to Perlmutter, languages with an obligatory subject in surface structure, such as English and French, are referred to as Type A languages. On the other hand, languages without an obligatory subject in the surface structure are Type B languages. Type B languages include: “Spanish, Italian, Serbo-Croatian, Arabic, Hebrew, Warlpiri and Basque”; Perlmutter insists that “every language must be either of one type or the other” (Perlmutter 1971:115). Most of the literature discusses null subject languages in comparison with non-null subject languages such as

English and French. It is understood that null subjects are permitted in languages such as Spanish and Italian because the referent of the implicit subject can be identified due to a rich morphological agreement system between the subject and the verb (Taraldsen 1978).

Following Taraldsen's (1978) insight that rich subject-verb agreement licenses null subjects, Rizzi (1982:142-143) formulated the null subject parameter. This parameter concerns whether or not an inflection can be pronominal and whether a pronominal inflection can be referential or non-referential. According to this parameter, null subject languages can be classified into two types: (a) expletive null subject language, in which the unexpressed subject can be an expletive; (b) referential null subject language, in which the unexpressed subjects are referential.

#### 4. 2. 2 Radical Pro-drop Languages

As seen in earlier studies, a common assumption about the possibility of null subjects is that null subjects are allowed in languages with a rich morphological verb-subject agreement (i.e. the null subject is recoverable from the rich morphology instantiated as Agr.) (Jaeggli and Safir 1989). In fact, the above proposals distinguish Spanish and Italian from non-pro-drop languages like English and French, but they cannot be extended to East Asian languages like Chinese, Japanese, and Korean, which are also known as null subject languages (or pro-drop languages). Those East Asian languages lack a rich system of morphological agreement but nevertheless license the omission of the subject pronoun (Huang 1984). Unlike Romance languages, occurrence of pro drop is not limited to subject position; pro drop can also occur in any other argument position, such as object. In the following example, (4.11), taken from our data, pro drop occurs not only in the subject position but also in the direct object position. From the

ambient discourse context, the unexpressed subject is interpreted as the speaker's brother, while the unexpressed direct object is considered to refer to English words or sentences.

(4.11)  $\emptyset$   $\emptyset$  *zenzen shira-n-katta.*

SUB DO at.all know-LNK-PAST

'[He] didn't know [them = English words/sentences] at all.'

(CJECC/O112/151)

The liberal distribution of zero pronouns is also known as “rampant pro drop” or “discourse pro drop.” Referring to “discourse-oriented languages” as defined by Tsao (1977), Huang (1984) proposed that null subjects are permitted because of the recoverability of topics in the wider discourse. In other words, the null subject can occur when the subject is a topic of the sentence and the topic is identified from the previous discourse as seen in example (4.11) above. More recently, Tomioka (2003) has proposed the Discourse Pro-drop Generalization to categorize East Asian languages as null subject languages:

(4.12) All languages which allow discourse pro drop allow (robust) bare NP arguments.

(Tomioka 2003:336)

### 4.2.3 Typological Approach to Null Subject Languages

On the basis of the discussion in the previous studies, Camacho (2013) and Roberts and Holmberg (2010) further extended the typology of null subject languages. Null subject languages can be classified into four types: (a) consistent null subject languages, (b) expletive null subject languages, (c) partial null subject languages, and (d) discourse pro-drop languages<sup>1</sup>. First,

<sup>1</sup> Camacho (2013) presents three types of null subject languages while Roberts and Holmberg (2010) presents the four types of null subject languages including expletive null subject languages. Expletive null subject languages such as German can be distinguished from non-null subject languages such as English.

*consistent null subject languages*, such as Italian, Spanish, Turkish, and Quechua, allow null subjects in all tense and all grammatical persons/numbers due to their rich subject-verb agreement system. When pronominal subjects are overtly expressed, they are generally used as a sign of topic change or contrastive interpretation. Second, *expletive null subject languages* allow expletive subjects to be null, but not referential ones. German, Dutch, Afrikaans, and a variety of creoles fall into this type. Third, in *partial null subject languages* including Finnish, Hebrew, Russian, and so on, null subjects can occur under restricted conditions associated with person, tense, and referentiality. For example, in Finnish and Hebrew, first and second person referential subject pronouns can appear as null in any finite clause, but not a third person pronoun. The opposite tendency is observed in Shipibo; first and second person subject pronouns are obligatorily expressed while third person subject pronouns are optionally null. Finally, Japanese, as well as Chinese and Korean, belong to the *discourse pro-drop* ('radical pro drop') languages. They are also referred to as topic-drop languages. In these languages, as we have seen in the previous section, unexpressed pronouns are allowed rather freely in any pronominal position. Null subjects are strongly associated with topics in discourse (Huang 1984).

Table 4.1: Summary of null subject languages

Type	Languages
Consistent null subject languages	Italian, Spanish, Turkish, Quechua
Expletive null subject languages	German, Dutch, Afrikaans, and a variety of creoles
Partial null subject languages	Finnish, Hebrew, Russian, Shipibo
Discourse pro-drop languages	Japanese, Chinese, Korean

To summarize, variation exists among null subject languages. As will be presented in § 4.5, a lot of the literature on null subject languages in a contact situation with a non-null subject language (e.g. English) has examined consistent null subject languages such as Spanish and Italian, which have a systematic verb-subject agreement system. To the best of my knowledge,

there is no study assessing the effects of contact of a non-pro-drop language such as English on Japanese, a discourse pro-drop language.

### 4.3 English as a Non-Pro-Drop Language

Although English is well-known as a non-pro-drop language, the literature has reported the occurrence of null subjects in Old and Middle English (Mitchell and Robinson 1986). In addition, Harvie (1998) has reported that the subject is omitted in English conversation as observed in the following example:

(4.13) Ø watch a movie or something, you know.

(source: Harvie, 1998:16)

Harvie employed a variationist approach and conducted a quantitative analysis to investigate how null subjects in English are patterning. It was found that null subjects are favoured in English if a clause containing a null subject is preceded by a clause which shares the same subject (i.e. in contexts of subject continuity). This finding indicates that an effect of subject continuity appears to be operative in both non-pro-drop and pro-drop languages.

Nariyama (2004) also reports the ellipsis of subjects in Australian English, using transcriptions of TV dramas, conversations, and casual letters. The majority of subject ellipsis in transcriptions occurs with the first person singular at 47.2% of the time, while ellipsis of a deictic pronoun 'it' was the most frequent in conversation (61.2%). In casual letters, the ellipsis of first person singular pronouns was found to be the most frequent. Nariyama (2003:258) presents the following synopsis of the major environments in which null subjects occur in the Australian data

she analysed: anaphoric deletion relying on linguistic context, deixis relying on situational context, dummy subject, and conventional expressions.

Furthermore, null subjects are often observed in English diaries, leading to so-called ‘diary drop’ (Haegeman 1990, 1997, 2002, 2013; Haegeman and Ihsane 1999, 2001; Scott 2010). In fact, diary drop is not limited to diaries; it is additionally observed in other written registers such as personal e-mails, texts, chats, telegrams, postcards, note-taking, and so on (Napoli 1982; Scott 2010; Weir 2009). Although it has been claimed that in English there are specific discourse constraints on the occurrence of null subjects in diaries (Haegeman 2013; Scott 2010; Weir 2008), researchers have also pointed out that there are syntactic constraints on the distribution of null subjects in English diaries (Haegeman 1990, 1997, 2013; Haegeman and Ihsane 1999, 2001). For example, null subjects in English diaries are observed only in declarative root clauses, as in the following example:

(4.14) Forgotten what I was supposed to do.

(source: Haegeman 1997:264)

Phonological constraints have also been implicated in the occurrence of null subjects in informal spoken English discourse. Weir (2009) claims that the first syllable in an intonational phrase, if unstressed, can be deleted in informal English, noting that the deleted element tends to be a subject pronoun. In addition, Weir compared nulls subjects in informal English with those in diaries. He suggests that his analysis of null subjects in informal English as a phonological phenomenon cannot be extended to written English, and that ‘diary drop’ is optimally analyzed as a syntactic phenomenon.

Summarizing, claims that English is a non-pro-drop language have been refined by a number of studies indicating that variable omission of subject pronouns in English is permitted in well-defined environments in certain written genres as well as in informal, spoken discourse. This omission appears to be constrained by discursive, syntactic, as well as possible phonological factors. Notwithstanding the recognition that English does allow null subjects in specific environments, there remains a general consensus of opinion that English is a non-pro-drop language as it does not license ‘rampant’ omission of subject pronouns characteristic of languages such as Japanese, nor does it instantiate the cluster of properties or “parametric effects” associated with consistent pro-drop languages such as Italian or Spanish (e.g. apparent violations of *that*-trace effects or a rich verbal morphology).

#### 4.4 Previous Studies on Null Subjects in Japanese

As we have seen in §4.2.2-§4.2.3, Japanese is classified as a discourse pro-drop language. Since nominals can drop not only in the subject position but also in any other position, such as direct object, indirect object, and so on, Japanese is also called a radical pro-drop language. It is reported that ellipsis of subjects is far more frequent than ellipsis of direct objects and indirect objects in conversation (Hinds 1983)<sup>2</sup>. Since Japanese lacks verbal agreement, as mentioned above, the unexpressed element is recoverable from the wider discourse context (Tsujimura 1996; Kayama 2003; Iwasaki 2013).

In Martin (1975:183-185), four types of null subject sentences in Japanese are discussed, as shown in the examples below. First, a specific referent of the subject can be easily identified as “it was probably ‘intended’ by the speaker and will be ‘understood’ by the hearer even though it

---

<sup>2</sup> Nariyama (2003) also reports the same tendency in written texts; subject ellipsis is more frequent than object ellipsis.

is not explicitly included in the message,” as in the example reproduced in (4.15). Martin labels this type, which has also been called “zero pronominalization,” as *optional ellipsis of the subject*.

- (4.15) (***Ame-ga***) *ya-n-da ka.*  
 rain-NOM stop-LNK-PAST Q  
 ‘Has [**the rain**] stopped?’

Second, “some sort of deictic reference is easily understood as the subject, even though the explicit verbalization of the reference is not easy to formulate” (Martin 1975:183-185). In the following example, (4.16a), the subject, according to Martin, can be understood as first person; however, it remains open whether the form is singular *watashi*, *watakushi*, *boku*, *atashi*, or plural *watashi-tachi*, and so on. In the same way, the subject in a sentence such as (4.16b) can be understood as second person, but the exact choice of form that could be overtly realized remains open.

- (4.16) a. *Kaeri-masu.*  
 go.home-DECL  
 ‘[I/We] am going home.’
- b. *O-kaeri desu ka?*  
 HON-go.home CPL Q  
 ‘Are [You] going home?’

Third, “expressions of time, weather, and other general conditions for which an arbitrary subject could be assumed” would not normally occur with any specific pronoun as the subject, as in the following examples, (4.17a-c) (Martin 1975:183-185).

(4.17) a. *Samui.*

cold  
‘[It]’s cold.’

b. *Osoku na-tta.*

late become-PAST  
‘[It] is getting late.’

c. *Yoroshii.*

good  
‘[It] is all good [to do it / not to do it].’

Lastly, the fourth type of sentence lacking an overt subject includes an animate subject with generic reference, such as *one, people, they, anybody*, and so on, as in the following examples in (4.18).

(4.18) a. *Ni ni ni-o tasu to, si ni naru.*

two to two-ACC add COND four to become  
‘Two plus two makes four.’ (Literally, ‘If [one/people] add two to two, [it] becomes four.’)

b. *Ningen to um-are-tara, hito no tame ni*  
human COMP born-PAS-COND others for.the.sake.of  
*tsukusanakereba naranai.*

must serve  
‘If [you] are born a human being, [you] must do all you can for the sake of others.’

A large number of studies have reported the phenomenon of subject omission in colloquial varieties of Japanese (Clancy 1980; Hinds 1978, 1982; Iwasaki 2013; Kameyama 1986; Kuno 1973a; Lee and Yonezawa 2008; Martin 1975; Nariyama, 2003; Okamoto 1985; Shibamoto 1983; Shibatani 1990; and others).

According to Shibatani (1990:360), ellipsis is more likely to occur in colloquial speech. For example, in (4.19) the speaker does not utter the first person singular pronoun (e.g. *watashi*)

to refer to himself since it is clear from the previous context that he is talking about how he came to Canada. Similarly, in (4.20) the speaker refers to her sister and her husband in the third person plural pronoun *karera* ‘they’ in the first sentence, but she drops the pronoun in the following sentence. The subject of the second sentence is co-referential with the subject of the first sentence: the speaker’s sister and her husband.

- (4.19)  $\emptyset$  *hitori de ki-ta.*  
 SUB one.person by come-PAST  
 ‘[I] came by myself.’

(CJECC/O106/135)

- (4.20) *Karera wa sugu kae-cha-tta-n-desu.*       $\emptyset$  *kyuu-ju-go nen gurai ni*  
 they TOP SOON return-AUX-PAST-LNK-COP      SUB ninty five years around in  
*nihon ni kae-cha-tte, mada nihon ni i-ru-n-desu.*  
 Japan to return-AUX-GER still Japan in existential-PRE-NM-COP  
 ‘They soon returned to Japan. [They] returned to Japan around 1995 and are still in  
 Japan.’

(CJECC/T013/118)

Occurrence of null subjects has been reported in the literature drawing on various data sources. Table 4.2 provides a summary of overall rates of null subjects in Japanese reported in diverse data sets. The oldest report is from the National Language Research Institute (hereafter, NLRI) in 1955: null subjects occur at 73.7% in conversation, at 37.3% in news and expository texts, and at 20% in novels<sup>3</sup>. In Clancy (1980), null subjects are observed at 73.2% in spoken narratives. According to other studies, rates of null subject use correlate with speaker sociodemographic characteristics. For example, Shibamoto (1983) reports null subjects occur at 73.3 % of the time in female speech and at 61.3% of the time in male speech in conversation (an average of 67.3%). Hinds (1983) has shown that differences in overall rates of null subject use appear to be genre-driven. For example, it was found that null subjects occur at 60.4% in

<sup>3</sup> NLRI (1955) has investigated the occurrence of null subjects in English and reported it occurs at 2%.

conversation and 73.9% in interviews while they occur at 47.5% in folktales<sup>4</sup>. Hinds also found that animate null subjects are more likely to occur than inanimate null subjects. Mizutani (1989) reports 77% null subjects, whereas Fry (2003) shows 69% null subjects in conversational data. Watanabe (1989) and Nariyama (2003) investigate null subjects in written narratives and texts. The former study reports 42.0% of null subjects in written narratives. The latter study has found that null subjects occur 56.3% of the time in written narratives, while they occur at 27.4% of the time in expository texts.

Table 4.2: Summary of rates of null subjects in various sources (spoken and written)<sup>5</sup>

Literature	Rate of null subjects	Data
NLRI (1955)	<b>73.7%</b>	<b>conversation</b>
	37.3%	news (expository texts)
	20.0%	novels
Clancy (1980)	<b>73.2%</b>	<b>spoken narratives</b>
Shibamoto (1983)	<b>73.3%</b>	<b>conversation (female speakers)</b>
	<b>61.3%</b>	<b>conversation (male speakers)</b>
Hinds (1983)	<b>73.9%</b>	<b>interviews (female)</b>
	<b>60.4%</b>	<b>conversation / interviews (male)</b>
	47.5%	a folktale
Mizutani (1985: 59)	<b>77.0%</b>	<b>conversation</b>
Watanabe (1989: 75)	42.0%	written narratives (textbook)
Nariyama (2000, 2003)	56.3%	written narratives
	27.4%	expository texts
Fry (2003)	<b>69%</b>	<b>conversation</b>

<sup>4</sup> Nariyama (2003) argues that the difference in the rate of ellipsis between interviews and conversations is due to gender difference. See Nariyama (2003) for more details.

<sup>5</sup> The rates in spoken data are bolded and highlighted.

Summarizing, null subjects tend to occur more frequently in colloquial speech (conversation and interviews) than in written texts. While most of the literature has reported rates of subject ellipsis of around 70% in conversation and interviews, the rate of null subjects in written Japanese is much lower around or below 50%<sup>6</sup>.

Shibamoto (1980, 1983) and Fry (2003) have examined null subjects in Japanese from a variationist sociolinguistic perspective. First, using data taken from sociolinguistic interviews, Shibamoto (1980) reported that female speakers are more likely to show null subjects than male speakers<sup>7</sup>. It has also been found that male speakers tend to show fewer null subjects in first person or second person deictic references than female speakers do. In addition, it was found that the referents of the null subject nominals among female speakers are more likely to be topic-related, indicating that female speakers tend to continue to talk about the same topic. However, this is not the case for male speakers. More recently, Fry (2003) has studied argument ellipsis using data extracted from telephone conversations<sup>8</sup>. Null subjects were found 69% of the time, which is more frequent than null direct objects. In addition, it was found that null subjects are more likely to occur with transitive verbs and when the referent of the subject is animate rather than inanimate, supporting previous claims by Hinds (1983). For social conditioning, Fry (2003) found no significant difference between male speakers and female speakers with regard to subject argument ellipsis.

---

<sup>6</sup> One of the reasons for this speech/writing difference may arise from the well-known distinction between oral and literate modes of communication: speech is typically highly context-bound while writing is much more decontextualized.

<sup>7</sup> However, no difference between female speakers and male speakers was found in other noun phrases such as direct object, indirect object, and so on.

<sup>8</sup> The data was taken from the CallHome Japanese (CHJ) corpus released by Linguistic Data Consortium in 1996. It consists of 120 spontaneous telephone conversations between native Japanese speakers (the caller in North America and the recipient in Japan).

Lastly, Fujii (1991) studied null subjects in Japanese from a historical perspective, considering the alleged influence of Western languages on Japanese. Fujii investigated diachronic change in the occurrence of null subjects using eight written texts composed from different periods. These texts are colloquial translations of a story called *Genji monogatari* ‘The Tale of Genge,’ which was originally written in the 11th century. It was found that subjects are more likely to be expressed overtly in the texts written between 1910 and 1940. Fujii claims that the increase in overt subjects during that period is due to the influence of Western languages because the period between 1910 and 1940 is considered as the time “when the influence of Western languages was strongly felt in many phrases of the language” (Fujii 1991:74). Fujii’s claims regarding the influence of Western languages on the increased rate of overt subject realization in Japanese rest on very speculative evidence: what this study is crucially lacking is a systematic comparison of patterns in Western languages that are believed to have had an impact on Japanese. As such, it is premature to conclude that the increased overt expression of subjects is the result of the influence of Western languages without first establishing whether the observed increase constitutes a *bona fide* change, and, if so, whether this change may instead be the product of internal tendencies rather than external influence.

In sum, null subjects in Japanese have been studied and discussed from various perspectives drawing on disparate datasets, methodological frameworks, and analytical techniques. Insofar as there is a consensus of opinion on null subjects in Japanese, there is general agreement that null subjects occur most frequently in colloquial speech varieties. However, a fundamental shortcoming in much previous research on null subjects in Japanese concerns the exaggerated importance attached to overall rates of variable occurrence. Rates are liable to fluctuate for all kinds of extralinguistic reasons related to genre, topic, and other

considerations (Torres Cacoullos and Travis 2010; Tagliamonte and Poplack 2001). As will be discussed in section 4.9 below, the present analysis transcends the conventional reliance on rates, and instead relies on the inspection of the linguistic conditioning of variable subject realization in order to scientifically pursue inferences of linguistic change in this sector of the grammar.

#### 4.5 Null Subjects in Contact Varieties

A number of studies have investigated the variable realization of subject pronouns in language contact settings (Bayley and Pease-Alvarez 1997; Flores-Ferrán 2004; Lapidus and Otheguy 2005; Montrul 2004, 2008; Nagy 2015; Nagy et al. 2011; Otheguy et al. 2007; Polinsky 1995, 1997; Polinsky and Kagan 2007; Sorace 2004; Sorace and Serratrice 2009; Silva-Corvalán 1994; Torres Cacollous and Travis 2010; Raña Risso 2010). Some of these studies present results that militate in favour of contact-induced effects on variable subject realization, while other studies report no effect.

It has been claimed that heritage language speakers exhibit different patterns of variable subject realization vis-à-vis monolinguals owing to the putative effects of contact, language attrition (or loss), or incomplete acquisition (Montrul 2004, 2008; Polinsky 1995a, 1995b, 1997, Polinsky and Kagan 2007; Sorace 2004; Sorace and Serratrice 2009). Language attrition or incomplete acquisition is claimed to be a consequence of sustained exposure to, or use of, a majority language, concomitant with reduced exposure to monolingual speakers of the L1. Attrition of the structure of the first language (L1) in contact situations with a second language (L2) is reported in numerous immigrant communities (see e.g. Seliger and Vigo 1991; Montrul 2004; Polinsky 1995). For example, Polinsky (1995:96) has reported the loss of pro drop among heritage speakers of Polish, Tamil, and Kabardian. In Montrul (2004), it was found that heritage

Spanish speakers of lower proficiency showed more overt subjects than monolingual Spanish speakers (and advanced heritage speakers). Toribio (2004:169-170) has also reported a possible effect of English on the use of overt subject pronouns in Spanish. Previous studies claiming a possible effect of English on the increased use of overt subject pronouns in heritage languages that license pro drop have frequently been conducted within an experimental paradigm. As Nagy (2015) points out, there are important methodological and analytical differences between studies of heritage language speakers conducted within an experimental framework and those that are articulated from a comparative variationist sociolinguistic perspective. Such differences may well account for discrepant findings across studies.

Nevertheless, even within the body of research based on a comparative variationist approach, a number of studies report divergent findings in relation to variable subject realization. Lapidus and Otheguy (2005) and Otheguy et al. (2007) report results that point to a putative contact effect on an increased rate of overt subject pronouns in heritage Spanish spoken in the USA. However, by and large, the weight of the available evidence marshalled by other scholars of Spanish-English bilingualism suggests that there has been no tangible effect of English on variable subject realization in US varieties of Spanish (see Bayley and Pease-Alvarez 1997; Flores-Ferrán 2004; Silva-Corvalán 1994; Torres Cacoullos and Travis 2010; Raña Risso 2010). None of these studies has confirmed the hypothesis that contact with English, a non-pro-drop language, increases the rate of overt pronouns in Spanish. Furthermore, they have found that the patterning of null subjects in heritage Spanish is similar to that of homeland monolingual varieties, militating against convergence with English.

More recently, Nagy et al. (2011) and Nagy (2015) have investigated the putative effect of contact with English on heritage languages such as Russian, Italian, and Cantonese spoken in

Toronto, Canada. Regarding Italian and Cantonese, no difference was found in the linguistic constraints operating on subject pronominal realization between speakers of the first generation and those of later generations, indicating that English had no effect on variant choice. By contrast, for speakers of heritage Russian, there were intergenerational differences in the constraint rankings for the internal factor groups of negation and grammatical person. However, such differences do not automatically imply English influence, as the linguistic conditioning of variability in negation and grammatical person in this heritage language differs from that of English (Nagy et al. 2011:140).

In sum, previous studies on the variable realization of subject pronouns in heritage language varieties in contact with English provide a strategic foundation for the present investigation embedded in a comparative variationist framework (Poplack and Tagliamonte 2001). As pointed out previously, the main diagnostic tools exploited in the current investigation for comparing linguistic structure across varieties are the constraint hierarchies generated by logistic regression analyses. We interpret these hierarchies as a window on the underlying grammar, and make use of the information they yield to illuminate structural affinities or divergence between linguistic varieties. In addition, the incorporation of social factors into the analysis enables us to ascertain which extralinguistic factors constrain the observed variability.

## 4.6 Variable Context

In this section, the methodological procedures adopted in the present study will be explained.

### 4.6.1 Envelope of Variation

The first important step in analyzing any variable linguistic system within a variationist framework is to identify and circumscribe the envelope of variation that determines the environments in which variation is observed (cf. Labov 1972). In the present study, the envelope of variation for variable realization of subject pronouns in Japanese is defined as those environments where either an overt subject pronoun or a null subject pronoun is able to occur. As is standard in variationist analysis (Tagliamonte 2006), this step necessarily entails the exclusion of contexts that do not meet the definition of the variable context, are anomalous or ambiguous, or are otherwise invariant.

### 4.6.2 Extraction Procedures and Exclusions

#### 4.6.2.1 Extraction

All tokens of verbs with an available subject slot filled by an overt pronoun were extracted. In addition, all tokens of verbs with an absent pronoun in the subject slot were also extracted. Verbs with an absent pronominal subject were required to have a “clearly ascertainable denotational subject” (Otheguy and Zentella 2013:51-53). Therefore, all and only verbs with an absent pronoun subject are included when their denotational subject is unambiguously recoverable. Since we focus on whether or not subject pronouns are overtly expressed, verbs with a full lexical subject were not considered in the study. Only verbs with an animate denotational subject were included in the present study; all tokens with an inanimate denotational

subject were excluded. We focus on the variable use of personal pronouns such as *watashi* ‘I,’ *anata* ‘you,’ *kare* ‘he,’ *kanojo* ‘she,’ their plural counterparts, and their variants, as illustrated below:

(4.21) First Person

- a. ***Watashi***-*wa* civilian *da-tta* *kara*.  
 I-TOP civilian COP-PAST because  
 ‘Because **I** was a civilian.’

(CJECC/O101/24)

- b. ***Atashi***-*wa* *sore* *deki-na-katta* *ne*.  
 I-TOP it can.do-NEG-PAST SFP  
 ‘I could not do it.’

(CJECC/O102/132)

- c. ***Boku***-*wa* *haaku-shi-te-na-i-desu* *ne*.  
 I-TOP grasp-do-GER-NEG-PRES-COP SFP  
 ‘**I** haven’t grasped it.’

(CJECC/O122/114)

- d. *Dakara* ***watashi-tachi*** *sugoi* *waruku* *yuw-are-mashita* *yone*.  
 therefore I-PL very badly say-PAS-DECL(PAST) SFP  
 ‘Therefore, **we** were very badly spoken about.’

(CJECC/O102/160)

- e. ***Boku-ra*** *nihon* *kara* *kai-masu kedo* *ne*.  
 I-PL Japan from buy-POL but SFP  
 ‘**We** buy them from Japan, though.’

(CJECC/O122/236)

- f. ***Wareware*** *mo* *eigo* *oboe-nakyanannai*.  
 we also English learn-AUX  
 ‘**We** also have to learn English.’

(CJECC/T012/102)

(4.22) Second Person

- Ontario no* *chizu* *de*, ***anata*** *mi-te*.  
 Ontario GEN map by you look-IMP  
 ‘**You** look at it, using the map of Ontario.’

(CJECC/O121/46)

(4.23) Third Person<sup>9</sup>

- a. ***Kare**-wa koko de soda-tta ko da kara...*  
 he-TOP here at grow.up-PAST child COP because  
 ‘Because **he** is the child who grew up here...’

(CJECC/T012/227)

- b. ***Kanojo**, sugoi kawaii to omo-u.*  
 she very cute COMP think-PRES  
 ‘(I) think that **she** is very cute.’

(CJECC/T010/101)

- c. ***Kare-ra**-wa supootsu-o ya-tte-na-katta node...*  
 he-PL-TOP sports-ACC do-GER-NEG-PAST because  
 ‘Because **they** were not doing sports,...’

(CJECC/O110/287)

Beginning at 10 minutes into each sociolinguistic interview, 80 tokens were extracted from each speaker to ensure that a balanced number of tokens were gathered from the sampling population. A total number of 1280 tokens were extracted from the corpus of heritage speakers and the same amount of tokens was extracted from the homeland benchmark variety, totalling 2560 tokens that were retained for analysis.

## 4.6.2.2 Exclusions

As indicated previously, any contexts that do not admit variation between overt and null subject pronouns were necessarily excluded from the analysis. A case in point regards conventionalized expressions such as various forms of greeting and set expressions (Nariyama 2000:22). For instance, Nariyama presents examples of *Sumanai* ‘sorry,’ *Arigatoo* ‘thank you,’

<sup>9</sup> One speaker used *sore* ‘it’ to refer to his friend instead of *kare* ‘he’ as in the following example. Those three tokens are included in the present study since the use of *sore* for *kare* may be a variant in the speaker’s dialect.

(1) *Sore-ga ee tomodachi ni na-tte ne.*  
 it-NOM good friend become-GER SFP  
 ‘He became a good friend.’

(CJECC/O112/135)

*Murimo nai* ‘(One) can’t help (it),’ and so on. Similar routinized expressions with null subjects are found in our data, as shown in (4.24-4.25). These conventionalized expressions are excluded from the present study.

(4.24) *Nanka hanashi-ga dondon zurete-tte, gomennasai ne.*  
 Somehow topic-NOM freely go.out.of.line-CT sorry SFP  
 ‘**Sorry** that the topic somehow goes out of line freely.’  
 (CJECC/T018/231)

(4.25) *Sorede, shibaraku shouganai kara, Ø amata shigoto sagashi-te*  
 and.then for.a.while cannot.help.(it) because SUB a.lot job look.for-CT  
*kondo Ø private ni hai-tta no.*  
 this time SUB to enter-PAST SFP  
 ‘And then, because **[I] couldn’t help it**, for a while [I] looked for jobs a lot, and then, [I] entered a private [company] this time.’  
 (CJECC/O101/96)

Secondly, examples such as those shown in (4.26-27) were excluded from the present study. The verbs *shaberu* ‘to chat’ and *modotte-kuru* ‘to return’ are nominalized and do not take an overt nominal argument as a subject. From a syntactic perspective, PRO is posited in the empty nominal argument position, as in these examples. Also, PRO can occur in the complement clause or the subordinate clause embedded in the matrix clause. For example, in (4.26) PRO is part of the complement of the main verbs *suki-da* ‘like-CPL.’ In (4.27), PRO is part of subject complement of the main verb *kanou-da* ‘possible-CPL.’ The reference of PRO is considered to be arbitrary. Shibatani (1990:361-362) labels it as PROARB. It is noted here that in Japanese the PRO occurs only in tenseless (non-finite) clauses (Shibatani 1990:285). Both clauses occurring with PRO in the examples are tenseless: *shaberu* and *modotte-kuru*. Also, it should be noted that the verb forms in these contexts function equivalently to the participle in English.

- (4.26) [PRO Shabe-ru-no] *suki-da kara.*  
 chat-PRES-NM like-COP because  
 ‘Because I like chatting.’

(CJECC/O106/238)

- (4.27) [PRO *Mata modo-tte-ku-ru* *koto wa*] *kanou da yo.*  
 again return-LK-COME-PRES NM(fact) TOP possible COP SFP  
 ‘Coming back (to Canada) again is possible.’

(CJECC/O103/9)

The subject of generalized animate referents, equivalent to English *one, they, people* is not expressed in Japanese (Kameyama 1985; Martin 1975:185; Shibatani 1990). In her conversational data, Shibamoto (1983:247) reports that 18.4% of ellipsed subjects are non-specific and thus interpreted as generic referents. In our data, when speakers describe general events or give instructions/explanations that require no specific referential subject, no overt subject is expressed in the sentence. The present study focuses on the verbs and their occurrence with specific subjects that are clearly recoverable from discourse. Since subjects encoding generic referents are reported in the literature to be ellipsed, every token with an identifiable generic referent was removed from the analysis.

In addition, lexical pronouns such as *minna* ‘everyone’ and reflexive pronouns such as *jibun* ‘self’ were not included in the analysis.

When the subject appears in postverbal position, as *karera* ‘they’ does in the following example where the speaker adds extra information after the null subject sentence, such tokens were excluded from the present study.

- (4.28) *Jibun de nantoka shi-te iki-te-ik-ou tte iu kimochi-ga*  
 oneself by somehow do-GER live-GER-go-VOL COMP say thought-NOM  
*a-ru yone, kare-ra-wa.*  
 have-PRES SFP he-PL-TOP  
 ‘**They** believe that they can manage to live somehow.’  
 (CJECC/T018/291)

Finally, and most importantly, if the grammatical subject is not recoverable from the ambient discourse context, then such tokens are excluded from the analysis as they could not be reliably coded. For example, a speaker uttered the following sentence, (4.29), when the interviewer asked him about his experience at work in Canada. However, the referent of the subject cannot be clearly recovered from the discourse context. It is unclear whether the unexpressed subject refers to ‘we,’ including the speaker, or ‘they,’ referring to his colleagues but not including the speaker himself. Accordingly, tokens with unrecoverable null subjects were not included in the study.

- (4.29)  $\emptyset$  *anmari kontakuto- $\emptyset$  tor-ana-i-wake-desu yo.*  
 SUB much contact-ACC take-NEG-PRES-SFP-COP SFP  
 ‘[We?/They?] do not make contact very much.’  
 (CJECC/O106/128)

## 4.7 Hypothesis

As we have seen earlier, some studies claim the effect of language contact with English are observed with regard to the variable realization of subject pronouns in minority languages (Polinsky 1995; Otheguy et al. 2007), whereas others find no contact-induced change (Flores-Ferrán 2004; Nagy et al. 2011; Torres Cacoullós and Travis 2010). Given that minority languages are often claimed to be susceptible to contact-induced change due to intensive contact with a majority language, it is hypothesized that the Japanese spoken by Japanese-English

bilingual speakers in Canada may exhibit an increase in the overt realization of subject pronouns compared to monolingual homeland Japanese speakers. We test this hypothesis by making use of the comparative variationist framework to affect a systematic comparison of variable subject pronoun realization in heritage Japanese in Canada with a commensurate non-contact vernacular benchmark, homeland Japanese, in order to determine whether there is empirical quantitative evidence of contact-induced change in the heritage variety.

## 4.8 Factor Groups

All tokens were coded according to a number of independent factors, or predictors, that have been hypothesized to have an effect on variant choice (i.e. the choice between an overt or null subject pronoun) in the language contact literature. All coded data are analyzed by means of GoldVarb X (Sankoff et al. 2005). The logistic regression procedure in this program enables the analyst to determine which independent variables make a significant contribution to variant choice when all of them are considered simultaneously. Social factors include speaker's length of stay in Canada for heritage Japanese speakers, age for monolingual Japanese speakers in homeland Japan, and sex for both cohorts. Linguistic factors to be examined include grammatical person, subject continuity (switch reference), previous realization of the subject pronoun (priming), clause type, and verbal transitivity. Each of these factors represents a testable hypothesis. These factors are described below.

### 4.8.1 Social Factors

First, the speaker's length of stay in Canada was examined for heritage Japanese to test the hypothesis that structural language change may be related to a long period of contact between

two languages (Thomason 2001; Thomason and Kaufman 1988; Winford 2003). In the homeland benchmark, speaker's age was included. The incorporation of an apparent time component in the homeland sampling population affords an opportunity to investigate whether there is any possible change in progress in this variety in relation to the variable realization of subject pronouns. Speaker's sex was also coded for, as sex differentiation in the variable expression of subject pronouns has previously been reported by Shibamoto (1980, 1983), who found that female speakers tend to omit subjects more often than male speakers do in Japanese.

## 4.8.2 Linguistic Factors

### 4.8.2.1 Grammatical Person

The first linguistic factor group to be examined is grammatical person. According to Kuno (1973a), in colloquial speech, third person pronouns are overtly realized considerably less often than first and second person pronouns<sup>10</sup>. Hinds (1978) also states that the use of the third person pronoun is rare, given the presupposition that a personal relationship exists between the referent and the speaker. Clancy (1980) also reports the infrequent use of third person pronouns in 20 Japanese film narratives. On the other hand, Martin (1975:332) states that first person pronoun *watashi* 'I' and especially the second person pronoun *anata* 'you' appear "a great deal less often," in comparison with their English counterparts, as direct pronominal reference is socially undesirable in Japan. In Japanese society, direct references tend to be avoided in order to be polite (Barke 2000:65). In such cases, the second person pronoun is usually replaced by a name and/or title (Martin 1975:332).

---

<sup>10</sup> However, Kuno (1973a:18) also states that third person pronouns are observed in the speech and writing of educated Japanese.

More recently, Lee and Yonezawa (2008) found that first person subjects and second person subjects were overwhelmingly omitted in face-to-face conversation, reporting rates of 84.5% and 88.5% for the null variants corresponding to the first and second persons respectively. Clancy (1980) compared Japanese and English with regard to the referential choice of third person human referents in narrative discourse. She found that in Japanese, ellipsis occurred 73.2% of the time, while a noun phrase accounted for the remaining 26.8% of contexts. No overt occurrence of third person pronouns was observed.

Table 4.3: Hypothesis for the effect of grammatical person on variant choice

Hypothesis	Favours Null	References
1 <sup>st</sup> pronoun	✓	Martin (1975); Shibamoto (1980,1983); Okamoto (1985); Kameyama (1986); Shibatani (1990); Nariyama (2003); Lee and Yonezawa (2008)
2 <sup>nd</sup> pronoun		
3 <sup>rd</sup> person pronoun	✓	Kuno (1973a); Clancy (1980); Hinds (1983)

The differential effect of grammatical person and number on the variable realization of subject pronouns has also been reported for other languages (Bayley and Pease-Alvarez 1997; Nagy 2015; Nagy et al. 2011; Otheguy et al. 2007).

Following previous work on both Japanese and non-Japanese languages, the present study also investigates grammatical person to see if any particular pronouns tend to be null in subject position. All tokens are coded according to grammatical person, as shown in the following examples.

(4.30) First Person

Atashi- $\emptyset$  *benkyou-shi-ta-koto-na-i.*  
 I-TOP study-do-PAST-fact-NEG-PRES.  
 ‘I have never studied.’

(CJECC/O101/38)

(4.31) Second Person

$\emptyset$  *nihon de hatara-ita-koto-ari-masu?*  
 SUB Japan in work-PAST-fact-have-DECL(PRES)  
 ‘Have [you] worked in Japan?’

(CJECC/O122/127)

(4.32) Third Person

a. *Sorede kare-wa shigoto-ga a-ru yone.*  
 And.then he-TOP job-NOM EXIS-PRES SFP  
 ‘And then, he had a job.’

(CJECC/O121/243)

b.  $\emptyset$  *nihongo-ga perapera sheb-ere-ru-n-de-su.*  
 SUB Japanese-NOM fluently talk-POT-PRES-NM-COP-PRES  
 ‘(He) can speak Japanese fluently.’

(CJECC/T004/58)

#### 4.8.2.2 Subject Continuity

Subject continuity has been widely investigated and reported as one of the strongest factors conditioning the variable realization of subject pronouns across languages such as Spanish, Italian, Russian, and Polish. In general, the null variant is favoured if the referent of the subject is the same as the subject of the preceding clause as in (4.33) (Silva-Corvalán, 1982; Cameron, 1993; Bailey and Pease-Alvarez, 1997; Nagy 2015; Nagy et al. 2011; Torres Cacoullos and Travis 2010). Conversely, Otheguy and Zentella (2012:60) report that subject pronouns appear overtly when there is a difference in reference from the subject in the previous clause (i.e.

‘switch reference’) as in (4.34). In the present study, the effect of switch reference is directly tested to see if variant choice is influenced by subject continuity<sup>11</sup>.

(4.33) Coreferential referent

*Kare-wa mou nihonjin-no-tomodachi-ga shocchuu ot-ta ne.*  
 He-TOP INTJ Japanese-GEN-friend-NOM always have-PAST SFP  
*∅ nihonjin to tsukia-tto-tta ne.*  
 SUB Japanese with associate-GEN-PAST SFP  
 ‘**He** always had Japanese friends. [**He**] was hanging out with Japanese people.’  
 (CJECC/O112/147)

(4.34) Different referent

*∅ soko kara imin shi-te-ki-ta hito de, de, kare-wa*  
 SUB there from immigration do-GER-COME-PAST person COP and **he**-TOP  
*koko de soda-tta ko da kara...*  
 here at grow up-PAST child COP because  
 ‘**[They]** emigrated from there, and **he** is the child who grew up here, so...’  
 (CJECC/T011/227)

<sup>11</sup> The literature on null subjects in Japanese discusses ‘switch reference markers,’ i.e. the conjunctive particles which indicate the identity of the subject in the preceding clause and the subsequent clause, as in the following examples: (sources: Iwasaki, 1993; Nariyama, 2002; Kameyama, 2003).

(1) SS (Same Subject) markers: *-nagara, -te-, -shi, -tsutsu, -∅, -tamani*

Example: *-nagara*

*∅ shucho o shi-nagara, ∅ chiimu waaku o tsuku-tte-ik-u-n-desu.*  
 SUB insistence ACC do-while SUB team work ACC make-CT-GO-PRES-SE-COP(PRES)  
 ‘**[They]** are doing teamwork while [**they**] are insisting on their opinion.’ (CJECC/T018/295)

DS (Different Subject) markers: *-to, -tara, -ga, -node, -yooni*

Example:

*Saikin wa iroiro ippai, ironna hito ga hai-tte-ki-ta-node, ma,*  
 recently TOP various many various person NOM enter-CT-come-PAST-because, ITJ  
*∅ nantomo i-e-na-i-desu kedo...*  
 SUB nothing say-POT-NEG-PRES-COP though  
 ‘Since various people came [to this area] a lot recently, [**I**] cannot say anything, though. (JECC/T010/219)

It has been reported that the distinction between SS markers and DS markers is not always an absolute indication of a switch reference (Iwasaki 1993). Although these markers were used for guidance in coding for the present study, coding was mainly done based on the interpretation of the context. Ambiguous tokens for subject continuity were excluded from the multivariate analysis.

Table 4.4: Hypothesis for the effect of subject continuity on variant choice

Hypothesis	Favours null subject	Favours overt subject	References
<b>Same referent</b> as the subject of the previous clause	✓		Silva-Corvalán (1982); Cameron (1993); Bailey and Pease-Alvarez (1997); Harvie (1998); Torres Cacoullos and Travis (2010); Nagy et al. (2011)
<b>Different (switched) referent</b> from the subject of the previous clause		✓	Ortheguy and Zentella (2012)

#### 4.8.2.3 Previous Realization (Priming)

Regarding the referential identification of null subjects, Kayama (2003:110) states that “when zero pronouns are used, they are identified by an overt nominal antecedent in the discourse.” It may be hypothesized that a null subject is likely to occur when the coreferential subject is expressed previously as the referent and can be identified easily as in the following example, (4.35). However, working with Spanish data, Torres Cacoullos and Travis (2010) found that a priming effect made a strong contribution to variant choice. The presence of a null subject in a preceding clause increased the likelihood that a null subject in a subsequent clause would be selected. Therefore, the present study investigates whether or not previous realization (priming) has an effect on variant choice in our data.

(4.35) Expressed previously

*Koukousei*                    *toka daigakusei-ga*                    *toshishita no ko-o*  
 high.school.student<sub>FP</sub>    university.student-NOM    junior                    GEN child-ACC  
*mendou mi-ru-n-desu.*    *De, sorede, Ø*    *okane-o kase-gu wake.*  
 care    look-PRES-NM-COP    then then                    SUB money-ACC    earn-PRES    SFP  
 ‘High school students and university students take care of juniors. Then, [they] make money.’

(CJECC/T018/278)

Table 4.5: Hypothesis for the effect of previous realization on variant choice

Hypothesis	Favours null subject	References
Expressed previously	✓	Kayama (2003)
Not expressed previously	✓	Torres Cacoullos and Travis (2010)

#### 4.8.2.4 Clause Type

According to Iwasaki (2013:280), null subjects occur more frequently in main clauses than in subordinate clauses in Japanese. As seen in (4.36), for example, the null variant is observed in the main clause whereas the overt variant, *watashi* ‘I,’ appears in the subordinate clause. On the other hand, the example in (4.37) contains the overt subject pronoun in the main clause and the null subject pronoun in the subordinate clause. Nagy et al. (2011) have also reported that subject pronouns appear null more frequently in the main clause in Cantonese and Russian in a contact situation with English. Accordingly, we examine the factor of clause type to see if similar effects can be found in our data.

- (4.36) *Watashi mo nihongo de hanashi-te-ta kara, ano,*  
 I also Japanese.language in speak-GER-PAST because uh  
 $\emptyset$  *wakari-masu.*  
 SUB understand-DECL(PRES)  
 ‘Because **I** was also speaking [to him] in Japanese, [**he**] understands (Japanese).’  
 (CJECC/O118/219)

- (4.37)  $\emptyset$  *shacho-ni-wa koe  $\emptyset$  kake-ta-koto-na-i-n-de ne,*  
 SUB president-DAT-TOP voice ACC apply-PAST-fact-NEG-PRES-SE-because SFP  
***boku wa sore-wa shiri-ma-sen kedo ne.***  
 SUB TOP it-TOP know-DECL-NEG though SFP  
 ‘Because [**I**] have never approached the president, **I** don’t know it.’  
 (CJECC/O122/122)

Table 4.6: Hypothesis for the effect of clause type on variant choice

Hypothesis	Favours null subject	References
Main Clause	✓	Iwasaki (2013); Nagy et al. (2011)

## 4.8.2.5 Verbal Transitivity

Fry (2003) reports that the subject of transitive verbs tends to be omitted in Japanese. For example, in his data, subject ellipsis occurs with the transitive verb *okuru* ‘to send’ at 89%, which is the highest rate, whereas it occurs with the non-transitive verb *nai* ‘not exist’ at the lowest rate of 23%. The present study tests the effect of predicate transitivity on the occurrence of null subject pronouns in the data to see whether or not the same effect reported by Fry (2003) obtains in the heritage and homeland Japanese data. Each token was coded according to whether it occurs in either a transitive predicate, as in (4.38), or a non-transitive predicate, as in (4.39). In our data, some transitive verbs are observed in passive forms, such as the transitive verb *yobu* ‘to call’ which appears in the passive form, *yobareru*, as in (4.40). Passive forms are coded separately from other transitive verbs.

## (4.38) Transitive predicates

- a.  $\emptyset$  *namae ga waka-n-nai-n-desu-kedo-ne.*  
 SUB name NOM know-GEN-NEG-GEN-COP-though-SFP  
 ‘[I] don’t know the name, though.’

(CJECC/T010/93)

- b. *Karera wa supootsu o ya-tte-(i)-na-katta-node,...*  
they TOP sports ACC do-CT-(PROG)-NEG-PAST-because  
 ‘Because they were not doing sports,...’

(CJECC/O110/287)

## (4.39) Intransitive predicates

- a. *Atashi  $\emptyset$  chotto koma-cha-tte ne.*  
 I NOM a little feel.at.a.loss-AUX SFP  
 ‘I felt a bit at a loss.’

(CJECC/O101/22)

- b. *Sorede  $\emptyset$  bosu to isshoni ocha ni i-ku no.*  
 Then SUB boss with together tea to go-PRES SFP  
 ‘Then, [we] go to tea with [our] boss.’

(CJECC/O101/45)

## (4.40) Passive form

*Demo ne, Ø daitai namae de yob-are-ru.*

but INT SUB generally first.name by call-PAS-PRES

‘But [I] am generally called by [my] first name.’

(CJECC/T010/230)

Table 4.7: Hypothesis for the effect of verbal transitivity on variant choice

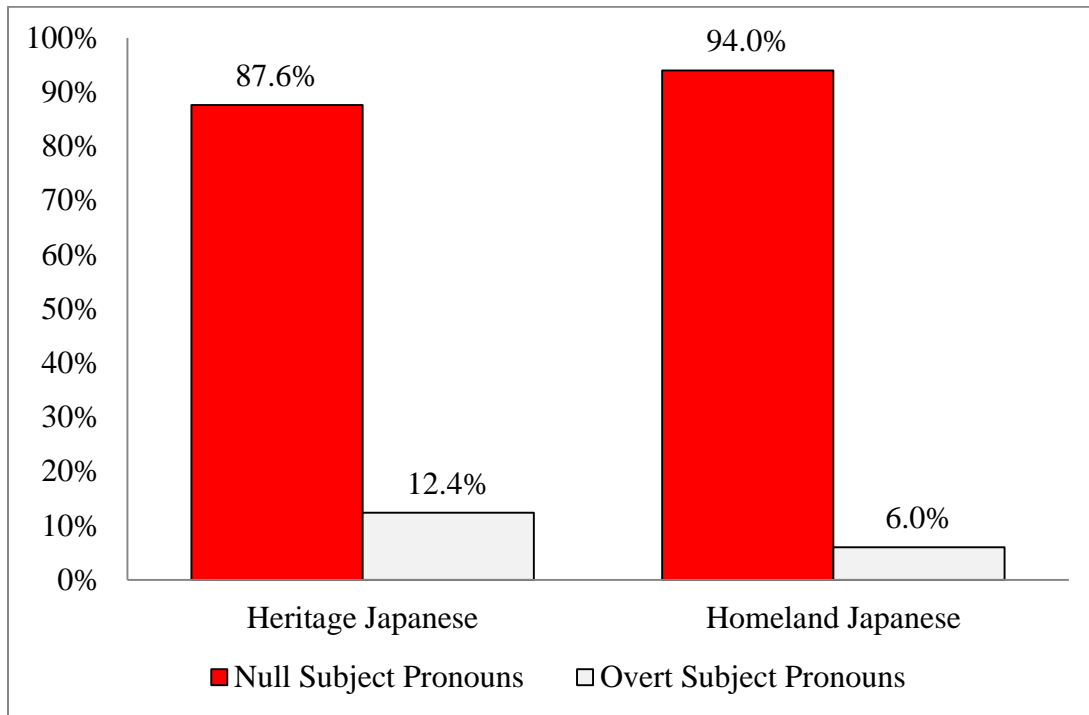
Hypothesis	Favours null subject	References
Transitive Predicates	✓	Fry (2003)
Non-transitive Predicates		

## 4.9 Results

### 4.9.1 Overall Distribution

Figure 1 presents the overall rates of variable realization of subject pronouns. In heritage Japanese, subject pronouns occur as null at 87.6% in the data. Subject pronouns occur as null at a rate of 94% in the homeland benchmark variety. Previous studies on null subjects in Japanese have reported the higher frequency of null subjects over overt subjects in conversation, as we already saw in §4.4: 77% in Mizutani (1985), 73.7% in NLRI (1955), 69% in Fry (2003), 67.3% in Shibamoto (1983), and 60.4% in Hinds (1983). Our results show much higher rates of null subjects; however, it should be noted that the previous studies on null subjects include not only pronouns, but also nominal subjects. Including the nominal subjects lowers the overall rates of null subjects. It is for this reason that the data in this study, including only overt pronouns and null pronouns, show higher rates of null realization than previous studies.

Figure 4.1: Overall rates of variable realization of subject pronouns



Our data show that null subject pronouns occur more frequently in homeland Japanese than in heritage Japanese. At first glance, then, the rate difference between heritage Japanese and the homeland baseline variety may suggest that the comparatively higher rate of overt subject pronouns in the heritage variety, double (12.4%) the proportion of overt pronouns in the homeland variety (6%), militates in favour of a contact-induced explanation. The rate differences between the two varieties are statistically significant, as assessed by a Chi-square test ( $p < .00001$ ).

However, rate differences are not, in and of themselves, reliable diagnostics of underlying grammatical differences between varieties. In order to ascertain whether there is divergence in the variable patterns underlying the two comparison varieties, we need to examine the fine-grained structure of the linguistic conditioning of variation. The results of this analysis are described in the following section.

### 4.9.2 Multivariate Analysis

By conducting a multivariate analysis, I examine how linguistic and social factors affect the realization of subject pronouns in Japanese when all of them are considered simultaneously. As explained in Chapter 3, we employ three lines of evidence (Poplack and Tagliamonte 2001) to interpret the results in order to discover underlying patterns in the data: statistical significance of factor groups (at the 0.05 level), relative strength (magnitude) factor groups as assessed by range value in each factor group, and the constraint hierarchy (the ordering of factor weights, or probability coefficients, within each factor group), which comprises favouring and disfavouring factors on variant choice. The comparison of constraint hierarchies will assess whether or not the same underlying grammar is shared by the two varieties of Japanese with regard to the variable realization of subject pronouns.

Table 4.8 presents the overall results of the multivariate analysis of social factors for the null subject pronoun in both heritage and homeland Japanese. In this table and all ensuing tables depicting a multivariate analysis, factors favouring a variant choice in the context are shown on the left side of the table. Numbers are interpreted as follows: numbers above 0.50 favour variant choice while numbers below 0.50 disfavour it. Numbers which hover around the median (i.e. 0.50) have no appreciable effect on the variant choice. Non-significant factor groups for the variant choice are indicated by the numbers in square brackets.

Table 4.8: Multivariate analysis of the contribution of social factors selected as significant to the probability of null subject pronouns; factor groups not selected as significant are in square brackets

	<b>HERITAGE JAPANESE</b>			<b>HOMELAND JAPANESE</b>		
Corrected mean:	.876			.940		
Total N:	1119/1280			1203/1280		
	FW	%	Token N	FW	%	Token N
<b>LENGTH OF STAY IN CANADA</b>						
Shorter	[.54]	89.4	572/640	N/A		
Longer	[.46]	85.8	549/640	N/A		
<i>Range</i>						
<b>AGE</b>						
Younger	N/A			[.50]	94.0	602/640
Older	N/A			[.50]	93.9	601/640
<i>Range</i>						
<b>SEX</b>						
Female	[.53]	89.1	570/640	[.50]	93.9	601/640
Male	[.47]	86.1	551/640	[.50]	94.1	602/640
<i>Range</i>						

As indicated by the corrected mean, the overall rate of null subjects is higher in homeland Japanese (.940) than in the heritage variety (.876). Inspection of the social conditioning of variation in the homeland Japanese variety reveals that the factor weights for age and sex hover at the median, indicating the absence of any effect.

From this information, it is inferred that variable subject realization is stable (i.e. not engaged in change) in the homeland variety. Turning to the heritage comparison variety, we see that length of stay in Canada is not selected as significant. In other words, the hypothesis that the longer speakers reside in Canada, the more likely they are to significantly decrease their use of the null variant, concomitantly increasing their rate of the overt variant, is not substantiated by this analysis. As sex effects are often interpreted as an important diagnostic of linguistic change in progress (Labov 1990), the absence of a significant effect offers additional evidence that no

change is underway. However, the most compelling evidence indicating that the underlying grammar in the comparison varieties is not divergent emerges from the linguistic conditioning of variability.

Table 4.9 presents the results of the multivariate analysis of linguistic factors for the null subject pronoun in both heritage and homeland Japanese. Inspection of the results in the homeland variety shows that grammatical person contributes the strongest effect to variant choice, as assessed by the range value of 53. Third person contexts strongly favour the null variant with a probability of .91, as reported in the literature (Clancy 1980; Hinds 1978; Kuno 1973a), while second and first person contexts have a disfavoured effect on variant choice with a probability of .47 and .38, respectively. The second ranked constraint is subject continuity with a relative strength of 30. In line with the variationist literature on pro-drop languages (Silva-Corvalán 1982; Cameron, 1993; Bailey and Pease-Alvarez 1997; Torres Cacoullos and Travis 2010; Nagy et al. 2011), the same referent favours the null variant with a probability of .63 whereas a different referent (i.e. switch-reference) disfavours with a probability of .33.

Table 4.9: Multivariate analysis of the contribution of linguistic factors selected as significant to the probability of null subject pronouns

	HERITAGE JAPANESE			HOMELAND JAPANESE		
Corrected mean:	.876			.940		
Total N:	1121/1280			1203/1280		
	FW	%	N	FW	%	N
<b>GRAMMATICAL PERSON</b>						
3rd person	.60	92.4	293/317	.91	99.6	223/224
2nd person	.69	93.2	41/44	.47	94.0	47/50
1st person	.46	85.6	787/919	.38	92.7	933/1006
<i>Range</i>	<b>23</b>			<b>53</b>		
<b>SUBJECT CONTINUITY</b>						
Same Referent	.62	92.4	621/672	.63	96.9	681/703
Different Referent	.36	80.1	432/539	.33	89.7	451/503
<i>Range</i>	<b>26</b>			<b>30</b>		
Non-significant linguistic factors: Previous Realization, Clause Type, and Verbal Transitivity.						

How does heritage Japanese compare with the homeland baseline variety? A first observation is that the same linguistic factors are selected as significant in the heritage variety by the regression analysis. However, their relative strength of effect differs from the homeland variety: subject continuity contributes the strongest effect, as assessed by the range value of 26, whereas grammatical person exerts a slightly weaker effect, as assessed by the range value of 23. Nevertheless, close inspection of the constraint hierarchy, our key heuristic for comparing the underlying grammars across varieties, reveals that the direction of effect for subject continuity is identical in the two comparison varieties; similarly to the homeland variety, in heritage Japanese, the null variant is favoured by the same referent in a preceding clause but disfavoured by a different referent, as indicated by the factor probabilities of .62 and .36, respectively. It is only within the grammatical person factor group that there is evidence of some modest structural divergence. In heritage Japanese, the null variant is favoured by second person and third person

with a probability of .69 and .60 respectively, whereas it is disfavoured with the first person, with a probability of .46. In the homeland variety, by contrast, the constraint hierarchy is reversed for the third and second person, but the first person still exerts a disfavoured effect on the choice of the null variant. To explore these patterns further, Table 4.10 presents a detailed breakdown of the distribution of overt and null variants by grammatical person.

Table 4.10: Overall rates of grammatical person in heritage Japanese and in the homeland benchmark variety

		HERITAGE JAPANESE				HOMELAND JAPANESE			
		Null	Overt	Total N	%	Null	Overt	Total N	%
1st person	N	787	132	919	71.8	933	73	1006	78.6
	%	85.8	14.4			92.7	7.3		
2nd person	N	41	3	44	3.4	47	3	50	3.9
	%	93.2	6.8			94.0	6.0		
3rd person	N	293	24	317	24.8	223	1	224	17.5
	%	92.4	7.6			99.6	0.4		

The rates of null realization in the second person are virtually identical in the comparison varieties: 93.2% in heritage Japanese and 94% in homeland Japanese. However, it should be noted that the proportion of the data accounted for by second person pronouns is less than 5% in both varieties (3.4% in heritage Japanese and 3.9% in homeland Japanese). Thus, we cannot rule out the possibility that the differences between the comparison varieties with respect to the behaviour of the second person may be an artefact of the paucity of data available for this grammatical person.

## 4.10 Discussion

What can we surmise from the results displayed in the previous section? When we compare the corrected means in heritage Japanese and the homeland benchmark variety, we observe that null subjects are more frequent in the latter variety. At first sight, this may seem to suggest that there is a difference in the quantitative patterning of null subjects in the two varieties. However, as Poplack and Tagliamonte (2001) point out, rates can fluctuate for all kinds of non-linguistic reasons and cannot be interpreted as diagnostics of differences in the underlying grammars of comparison varieties. In addition, the ranking of the two significant factor groups, grammatical person and subject continuity, is in a different order, supporting the idea that there is a divergence between the comparison varieties of Japanese. However, a more accurate heuristic for comparing underlying grammars is the constraint hierarchy (Poplack and Tagliamonte 2001:93-94), construed as a window on a portion of the underlying grammar. The comparison of constraint hierarchies in the heritage and homeland varieties reveals that, by and large, they show the same direction of effect for all independent variables operationalized in the study. It suggests that variable grammar is shared by the comparison Japanese varieties<sup>12</sup>. The

---

<sup>12</sup> Previous studies examining the linguistic conditions for English pro drop have found that null subjects in English are favoured by the same referent as the preceding subject (Harvie 1998; Nagy et al. 2011; Torre Cacoullous and Travis 2010). Subject pronoun realization in heritage Japanese data is conditioned in the same way as English with regard to the subject continuity factor group. However, it cannot simply lead us to claim that this is due to contact with English; first, an identical constraint hierarchy is observed between heritage Japanese and homeland Japanese in the present study as seen in Table 4.9. Second, the same direction of effect has been reported in other languages such as Spanish, Italian, Cantonese, Russian, and so on, indicating a universal linguistic feature. According to Poplack and Tagliamonte (2001:101), “the conflict site is typically instantiated as a difference in constraint hierarchy.” However, the same direction of effect is found between heritage Japanese and English as observed in the literature. Therefore, this would not be a good diagnostic for assessing contact-induced change since this kind of effect, whereby a null subject is favoured by the same referent as the preceding subject, seems to be common in most languages. In addition, although our results show that grammatical person conditions subject expression, it seemed to have no effect on variant choice in previous studies on English pro drop. According to Harvie (1998), null subjects occur in third person contexts at 35%, in second person contexts at 34%, and in first person contexts at 29%. As shown in Table 4.10, our results of heritage Japanese show a different tendency. Summarizing, convergence of heritage Japanese spoken in Canada toward English cannot be confirmed with regard to the variable realization of subject pronouns.

only anomalies reside in the quantitative patterning of grammatical person.

Consideration of the multiple lines of evidence revealed by the multivariate analysis indicates that there is little compelling evidence to suggest that the grammar underlying variable subject pronominal expression in heritage Japanese diverges markedly from the underlying grammar instantiated by the homeland benchmark variety. Viewed in the aggregate, these results converge in demonstrating that there is no substantial evidence of contact-induced change in this sector of the grammar.

# CHAPTER 5

## VARIABLE CASE MARKING ON NOUNS

### IN SUBJECT AND DIRECT OBJECT POSITION

#### 5.1 Introduction

The next linguistic variables targeted for the investigation of contact-induced change in heritage Japanese are: (a) variable *case marking of subject nouns* and (b) variable *case marking of direct object nouns*. In Japanese, functional relations are expressed by postpositional particles (Johnson 2008; Shibatani 1990; Tsujimura 1996). Particles are used for indicating the case relation of nominals: nominals in an argument position are marked by the relevant particle for case marking, such as a nominative, accusative, and so on. From a prescriptive point of view, the nominative case marking particle *ga* should appear obligatorily for nouns in subject position. Similarly, the accusative case particle *o* is also obligatorily prescribed for nouns in object position. In (5.1), which is repeated from (1.2a), the subject of the sentence *guntai* ‘military force’ is marked by the nominative case particle *ga* and the object *toukaku* ‘preeminence’ is marked by the accusative case particle *o*.

- (5.1) *Horede,      guntai      ga      bochibochi ne,      toukaku      o*  
 then            military.force NOM      gradually      INT      preeminence ACC  
*arawashi-te-ki-ta.*  
 show-CT-COME-PAST  
 ‘Then, the military force has gradually risen to preeminence [in Japan].’  
(CJECC/017a/203)

Although prescriptive norms state that nouns in subject and direct object argument position should be followed by the relevant particles for case marking, it has been reported in the literature that these particles can be dropped in colloquial speech (Martin 1975; Murasugi and Sugisaki 2008; Shibatani 1990; Tsutsui 1984; Matsuda 1995; NLRI<sup>1</sup> 1995, and others). This is precisely what is observed in the corpus of heritage Japanese speakers. In example (5.2), the noun *rikugun* ‘military’ omits the case particle for marking the subject while the noun *keni* ‘authority’ is marked by the accusative case marker *o*. In contrast, in (5.3), which is repeated from (1.2b), the nouns *hito* ‘person’ and *inu* ‘dog’ omit the nominative marker *ga* for marking the subject as well as the accusative case particle *o* for marking the direct object, respectively.

- (5.2) *Houshite rikugun Ø zettaiteki keni o nigichatta.*  
 And.then military NOM absolute authority ACC seize-PAST  
 ‘And then, the army took over.’  
 (CJECC/017a/208)

- (5.3) *Yoku sono hito Ø ne, chichana inu Ø tsure-te sanpo-shi-te-ta*  
 often that person NOM INT small dog ACC take-CT walk-do-CT-PAST  
 ‘That person often took a walk with a small dog.’  
 (CJECC/010/281)

The main objective in this chapter is to investigate the hypothesis of contact-induced change with respect to the variable case marking on subject nouns and direct object nouns. Adopting the same approach used in the previous chapter in the investigation of the variable realization of subject pronouns, we will assess whether the grammatical structure of heritage Japanese is influenced by English by comparing the patterns of heritage Japanese with those of the homeland benchmark variety and employing an empirically accountable quantitative approach. The potential influence of English on the structure of heritage Japanese is motivated by the widespread assumption that contact is believed to accelerate the attrition or loss of

<sup>1</sup> NLRI stands for the National Language Research Institution, as introduced in Chapter 4.

features (e.g. case marking) in one language that do not have a counterpart in the majority language (Poplack and Levey 2010:393). It is hypothesized that if contact-induced change exists, then its effects should emerge from overall rate differences in case marking when heritage and homeland Japanese are compared, and, most important of all, from differences in the grammar underlying the observed variability, as revealed by a systematic comparison of constraint hierarchies in the two varieties of Japanese.

The remainder of this chapter is organized as follows: the literature pertaining to case marking features is reviewed in §5.2-§5.3. Section 5.4 then describes the methodological procedures adopted in the present study to circumscribe the envelope of variation and extract the target variable. Section 5.5 presents the hypothesis for the investigation. In §5.6, social and linguistic factors operationalized in the study are discussed in order to explore potential effects of contact-induced change. Section 5.7 contains the results of the investigation, and this is followed in Section 5.8 by a discussion of the major findings with a view to examining their pertinence to the research objectives presented at the outset of this chapter.

## 5.2 Case Marking in Japanese

### 5.2.1 Theoretical Approach

In the 1960s and 1970s, case marking in Japanese was approached from the perspective of generative grammar (Kuroda 1965; Kuno 1973a; Tamori 1977). Following Chomsky's transformational grammar (1965), it is assumed that some particles are originally in deep structure while others are introduced by a set of transformational rules in surface structure (Kuroda 1965). The latter particles include case-marking particles such as the subject case marker *ga*, the direct object case markers *o* and *ga*, the indirect object case marker *ni*, and the



importantly, Tamori found a syntactic constraint conditioning variable case marking: deletion of case markers does not occur after scrambling constituents in a sentence if a meaning change occurs. He provides the following examples.

- (5.5) a. *Taroo-ga Hanako-o aisi-ta.*  
           NOM          ACC          love-Past  
       ‘Taro loved Hanako.’  
   b. *Taroo  $\phi$  Hanak  $\phi$  aisi-ta.*  
   c. *Hanako o Taroo ga aisi-ta.*  
   d. *Hanako  $\phi$  Taroo  $\phi$  aisi-ta.*

(source: Tamori 1977:258 with modified gloss.)

According to Tamori, the examples in (5.5a-c) produce the same meaning even though the case markers are deleted in (5.5b) and the subject *Taroo-ga* and the direct object *Hanako-o* are scrambled in (5.5c)<sup>3</sup>. In (5.5d), both scrambling and deletion of case markers are observed; in this case, it is ambiguous whether or not the subject and the direct object are identified in the same way as (5.5a). This sentence may even change the interpretation to *Hanako* loved *Taroo*, which is not equivalent to (5.5a). It is therefore claimed that deletion of case markers is not possible in a scrambled structure if this entails a change in meaning. Finally, Tamori proposes a hierarchy of zero-marking particles as follows:

- (5.6) Subject Marker > Direct Object Marker > Indirect Object Marker > Other Particles

(source: Tamori 1977:257)

According to the hierarchy, the subject marking particle *ga* is most likely to be deleted, followed by the direct object marking particle *o*, the indirect object marking particle *ni*, and then other particles (e.g. the interrogative marking particle *ka*).

<sup>3</sup> Tamori (1977:258) mentions that the simultaneous deletion of both case markers, *ga* and *o*, is questionable in actual speech. However, it is still possible to identify the subject and the direct object and arrive at the same meaning as (5.4a).

It is generally claimed that because the nominative case marker *ga* is never truly deleted, and any deleted particle marking the subject NP would have to be the topic marker *wa* (Kuno1973a, 1973b). In other words, *ga* is obligatory for the subject while *wa* is optional (1973b:223-224). Saito (1985) approaches this generalization from the perspective of government and binding theory. He claims that nominative case and accusative case follow the Case Filter proposed by Chomsky (1981:49).

(5.7) Case Filter

\*NP if NP has phonetic content and has no Case.

According to Saito, abstract accusative case is assigned by the verb regardless of the presence or absence of the accusative case marker *o* in Japanese. It is suggested that “the accusative Case marker *o* is in some sense a phonetic realization of the abstract object Case” (Saito 1985:206). On the other hand, nominative case is not assigned by any element but exists inherently in its subject. If the nominative case marker *ga* is deleted, it indicates a violation of the Case Filter because the subject conveys no case. Therefore, a null-marked subject is considered to represent the deletion of the topic marking particle *wa*, supporting Kuno’s (1973a, 1973b) claim.

Since Tamori (1977), more studies have discussed the constraints on ellipsis of case-marking particles. For example, a syntactic constraint, verbal adjacency, is reported by Tsutsui (1984) and Saito (1985). It is claimed that the ellipsis of case markers is more natural if the subject and the direct object are immediately followed by the predicate. Furthermore, Tsutsui (1984) proposes general rules for particle ellipsis and different rules for specific particle ellipsis (i.e. nominative case *ga* and accusative case *o*) in conversation. Those rules cover various linguistic levels such as syntax, phonology, pragmatics, and style of the conversation. For

example, as one of the general rules, Tsutsui (1984:90) discusses formality: “[t]he lower the formality level is, the more natural the ellipsis of case particle is.” This is one of the first formulations regarding the putative effect of formality on the ellipsis of case markers. Masunaga (1988) discusses pragmatic constraints on ellipsis of case markers. She claims that ellipsis of case markers is acceptable when the NP is de-emphasized by the presence of sentence-final particles. Further details are presented in § 5.6.2.3.

### 5.2.2 Empirical Studies

Although case-marking particles are normatively prescribed for nominals in subject and direct object position, a number of studies acknowledge that these particles can be variably omitted in speech (Martin 1975; Murasugi and Sugisaki 2008; Shibatani 1990; Tsutsui, 1984; Matsuda 1995; NLRI 1995, and others). According to the research on subjects by NLRI (1955:113), the subject of a sentence occurs without a case-marking particle at 29% in colloquial speech whereas the rate of case-marking omission on subjects is lower in news broadcasting and newspapers, 0% and 9.7%, respectively. The NLRI study confirms that subject nouns are followed by a particle in written newspapers, which are considered to reflect prescriptive grammar, although speech is much more prone to case-marking omission. In addition, it has been observed that ellipsis of a particle occurs more frequently in colloquial speech than in formal speech such as broadcasting news. Furthermore, it was reported that in colloquial speech, where a subject occurs without a particle at an overall rate of 29%, the subject occurs with the nominative case marker *ga* at 33.5%, with the topic marker *wa* at 23.7%, with the particle *mo* ‘too, also’ at 6.7%, and with others such as *tte*, *datte*, and *demo* at 7.3%.

More empirical studies have been conducted since the mid-1980s. Shibamoto (1985) carried out sociolinguistic research on various linguistic variables including particle ellipsis, focusing on gender differences. The data contain natural speech collected from speakers in Tokyo. Regarding the subject case markers, it was found that female speakers are likely to use zero markers more frequently than male speakers (23.9% vs. 11%)<sup>4</sup>. As for the direct object case markers, the difference between female speakers and male speakers was even greater: zero marking is used much more often by female speakers than male speakers (40.4% vs. 25.3%). Since Shibamoto did not analyze ellipsis of case markers in terms of linguistic constraints, her findings with regard to gender differences she reported remain inconclusive. However, her study is among the first to explore social constraints on the variable omission of case markers.

A growing body of research has investigated case-marking ellipsis, drawing on quantitative techniques and using natural speech data. Two studies in particular, Takano (1998) and Matsuda (1995), approached variable case marking within a variationist framework. Takano (1998) investigated gender differences with regard to the ellipsis of the nominative case marking particle *ga* and the topic marking particle *wa* in the subject position. First, a multivariate analysis revealed that particle ellipsis is intricately conditioned by hierarchically ordered linguistic constraints. Secondly, although Shibamoto (1985) simply observed that the ellipsis of particles is used by female speakers more than by men, Takano found that the frequency of the ellipsis varies according to the different types of participant frameworks: single-sex dyadic, mixed-sex dyadic, and heterosexual group. There was a gender-related difference in single-sex dyadic conversations: ellipsis occurs more frequently in female-to-female conversation than in male-to-

---

<sup>4</sup> Shibamoto (1988) classified case markers into the following groups: subject marker *ga* or object markers *ga/o*, and focus markers *mo* and *wa*, case-marking substitutes such as *tte* and *nante*, and zero marker.

male conversation. However, no significant difference was observed regarding the frequency of the ellipsis in mixed-sex dyadic and in the heterosexual group.

Matsuda (1995) conducted a large-scale investigation of the variable accusative case marker *o*. Both linguistic and social factors were exhaustively examined using natural speech data from Tokyo Japanese. Through a multivariate analysis the study identified significant internal factors and social factors conditioning the speaker's variant choice. Matsuda (1995) reported that the best model to describe the variation in Tokyo Japanese is by incorporating factors such as the nature of the NP, word order, speaker's area of residence, and style. First, he found that zero marking occurs most frequently with *wh*-pronouns and least frequently with clausal NPs, as schematized in the following hierarchy:

(5.8) Zero-Marking Hierarchy in Tokyo Japanese

*wh*-pronoun > NP > non-*wh*-pronoun > clausal NP

(source: Matsuda 1995:168)

Furthermore, zero marking was favoured where the object noun is adjacent to the verb. As for speaker's sex, female speakers are more likely to show zero marking than males. He also found a correlation between zero marking and residential area: zero marking is favoured more by speakers from the downtown area of Tokyo than those from the uptown area. Finally, it was observed that zero marking occurs more frequently in casual speech, rather than careful speech, supporting Shibamoto's (1985) finding.

In sum, both Takano (1998) and Matsuda (1995) demonstrated that the variable grammar underlying case marking can be revealed by detailed quantitative examination of the multiple factors that condition variant choice. Using a similar methodology, the present study will attempt

to elucidate the social and linguistic conditioning of case-marking variability for subject and direct object nouns in heritage and homeland Japanese.

### 5.3 Case Marking in Contact Varieties

A number of studies have investigated case loss or case reduction in language contact settings (Anderson-Martine 1976; Clyne 2003; Huffines 1989, 1991; Larmouth 1974; Meyerstein 1969; Preston 1986; Polinsky 1995, 1997). Heritage language speakers are claimed to exhibit different patterns of case marking vis-à-vis monolingual speakers, due to the putative effects of language attrition or incomplete acquisition (Montrul 2004, 2008; Polinsky 1995, 1997, 2006; Polinsky and Kagan 2007; Sorace 2004; Sorace and Serratrice 2009).

Larmouth (1974) found change in the case system in Finnish spoken in Minnesota, United States, via comparisons of four generations. Elision of case endings was found to be more likely to occur in later generations “as a result of interference from English” (Larmouth 1974:360). A loss of the inflectional system has been reported in various languages in contact with English: for example, Polish in Preston (1986) and Pennsylvania German in Huffines (1989, 1991). Preston (1986:1020-1021) found loss of case marking in Polish spoken in New York: heritage Polish speakers showed loss of dative, instrumental, and locative case for adjectives as well as loss of dative case for nouns. Huffines (1989, 1991) discussed language change and convergence as processes of attrition in Pennsylvania German, based on her finding that the dative case and the accusative case are merging, indicating an ongoing loss of dative case marking. Huffines (1989) claims that this is the result of contact-induced change due to English influence.

A number of experimental studies on heritage languages in the United State have reported loss or reduction of case in contact languages. Meyerstein (1969) analyzed a list of

prepositional phrases in the text produced by immigrant Slovak speakers and found that the second generation of heritage Slovakian speakers was showing a reduction to a three-case paradigm in case marking, compared with the first generation of speakers. In Polinsky (1997:375-382), heritage Russian speakers have reportedly lost the case system, which is preserved by monolingual speakers. For example, instrumental case was lost from the context where either nominative or instrumental case is supposed to be available. Also, the lexically-governed genitive case was observed to be replaced by the accusative case or by the nominative case, depending on the speakers' proficiency in Russian. This is considered to be an outcome of language attrition. Laleko and Kawamura (2011) reported that heritage Japanese speakers show different patterns of case marking and topic marker use when compared to patterns of monolingual speakers, including the overuse of the nominative case marker *ga* and the underuse of the topic marker *wa*.

Building on previous studies of case marking in heritage language varieties in contact with English, the present study focuses on variable case marking (i.e. the presence or absence of a case marker) in heritage Japanese spoken in Canada. In order to transcend the methodological shortcomings and data limitations in previous research on this topic, the present study uses spontaneous speech data to analyze case marking quantitatively and systematically in heritage Japanese, drawing on the infrastructure of comparative variationist sociolinguistics (Poplack and Tagliamonte 2001). We emphasize the importance of constraint hierarchies generated by a logistic regression analysis to interpret the underlying grammar conditioning variable case marking in heritage Japanese and homeland Japanese and to evaluate similarities and dissimilarities between these varieties, with a view to ascertaining whether heritage Japanese

diverges from the homeland benchmark as a possible result of extended contact with English. In the following section, the methodology used in this study will be described.

## 5.4 Variable Context

This section describes the methodological procedures adopted in the present study to investigate variable case marking in heritage Japanese.

### 5.4.1 Envelope of Variation

In Japanese, nouns are case marked by postpositional particles. English does not have such a property. For example, according to normative accounts, the nominative case-marking particle *ga* should appear obligatorily for nouns in subject position. Similarly, the accusative case particle *o* is also obligatorily prescribed for nouns in object position. In (5.9), the subject *ooya-san* ‘landlord’ is marked by the nominative case particle *ga* while the direct object *inu* ‘dog’ is marked by the accusative case particle *o*.

- (5.9) *Ooya-san* **ga** *sugoku ii hito de saikin inu* **o**  
 landlord NOM very good person COP.and recently dog ACC  
*kai-hajime-ta no.*  
 have-start-PAST SFP  
 ‘My landlord is a very good person, and he got a dog recently.’

(CJECC/T003/22:39)

In the present study, the envelope of variation in examining case marking on nouns includes all and only nouns occurring in subject position and direct object position. In accordance with the principle of accountability (Labov 1972), we include not only case-marked nouns but also null-marked nouns. Nouns in contexts where no variation is found are excluded from the analysis, as detailed in the following section.

## 5.4.2 Extraction Procedures and Exclusion

### 5.4.2.1 Extraction

As described in 1.5.2, the present study compares case-marking patterns of Japanese nouns and English-origin nouns in heritage Japanese with those of Japanese nouns and loanwords in the monolingual benchmark variety in Japan. There are many loanwords from various source languages such as Chinese, Dutch, Portuguese, and Sanskrit in modern Japanese (Shibatani 1990; Stanlaw 2004). In the present study, however, loanwords are defined as English-origin loanwords to draw comparisons with English-origin nouns in heritage Japanese in contact with English.

All tokens of nouns occurring with the nominative case marker *ga* in subject position and with the accusative case marker *o* in direct object position were extracted, while all tokens of nouns without a case marker were also extracted. Subjects and objects in sentences could also be marked by a particle other than *ga* or *o*. When a noun is the subject and also the topic in a sentence, the nominative case marker *ga* can be replaced by the so-called topic marker *wa*. In (5.10), the subject noun *hito* ‘person’ is marked by the topic marker *wa*<sup>5</sup>. In the same way, when an object noun is the topic in a sentence, the accusative case marker *o* can be replaced by the topic marker *wa* as in (5.11). The topic marker *wa* can be omitted in colloquial speech as well as in formal speech (Kuno 1973a, 1973b; Martin 1975).

---

<sup>5</sup> The topic marker is used as a contrastive marker (c.f. Iwasaki 2003):

(i) *taro-wa gakusei da ga, mami-wa gakusei ja-nai.*  
 (name)-TOP student COP but (name)-TOP student COP-NEG  
 ‘Taro is a student, but Mami is not a student.’ (source: Iwasaki 2013:245)

The same noun phrase can be both topicalized and contain contrastive elements (Iwasaki 2003:245).

- (5.10) *Ano hito wa kokode yatow-are-ta-n-desu ne.*  
 that person TOP here hire-PASS-PAST-SE-COP(PRES) SFP  
 ‘That person was hired here.’

(CJECC/O106/184)

- (5.11) *Denwa wa yoku su-ru-kara ne.*  
 phone TOP often do-PRES-because SFP  
 ‘Because [I] often make a phone call.’

(CJECC/O118/54:02)

Although the nominative case particle *ga* usually marks the subject in a sentence, the literature has reported that in certain constructions, an object can also be marked by the nominative case particle *ga*, instead of the accusative case marker *o* (Kishimoto 2000; Koizumi 2008; Kuno 1973; Martin 1975). When an object is marked by the nominative case particle *ga*, this is called the nominative object construction (Koizumi 2008). For example, the nominative case particle *ga* is used to mark the object of stative transitive verbs such as *wakaru* ‘understand, sense’ in (5.12), *dekiru* ‘can do’ in (5.13), and so on. When an action verb occurs with the auxiliary adjective morpheme *-tai* ‘want,’ the object of the predicate is also marked by the nominative case particle *ga* as in (5.14). Therefore, we extract all nouns and objects nouns taking case particles *ga*, *o*, and *wa*, as well as nouns that omit those case markers, in order to account for variability.

- (5.12) *Kaze-no nagare ga waka-ru.*  
 wind-GEN flow NOM sense-PRES  
 ‘[I] sense the flow of wind.’

(CJECC/010/298)

- (5.13) *Betu-no kaigai-ryokou ga deki-na-kute.*  
 another -GEN abroad-travel NOM can.do-NEG-GER  
 ‘I cannot make another trip abroad.’

(CJECC/003/106)

- (5.14) *Rekishi-no-isan* *ga* *mi-tai.*  
 history -GEN -heritage NOM see-want  
 ‘I want to see items of historical heritage.’

(CJECC/003/97)

Beginning at 10 minutes into each sociolinguistic interview, up to 25 relevant tokens for both subject position and direct object position respectively in each discourse (i.e. English-origin nouns and Japanese nouns in heritage Japanese, and Japanese nouns and loanwords in the homeland benchmark variety) were extracted. This yielded an initial total of 1132 tokens of subject nouns and 1183 tokens of direct object nouns.

#### 5.4.2.2 Exclusions

##### 5.4.2.2.1 Categorical Environments

As explained in the previous chapter on the variable realization of subject pronouns, any categorical environment in which only one form (i.e. either case marking or null marking) is possible are not be included in the analysis of case marking as such contexts are invariant. This is the case for the particle *mo*. When case particles of nouns in subject position and object position are replaced by the particle *mo*, meanings such as *also*, *too*, *as well as* and so on, are added to a sentence. As shown in example of (5.15), the subject noun *koshi* ‘lower back’ is followed by the particle *mo*. The particle *mo* cannot be omitted without an alternation in the referential meaning of the sentence. Since the present study focuses on overt/null marking, nouns followed by the particle *mo* are not included.

- (5.15) *Koshi* *mo* *ita-ku-na-tta.*  
 lower.back also sore-CONJ-become-PAST  
 ‘The lower back also became sore.’

(JECC/O101/291)

## 5.4.2.2.2 Proper Nouns

All proper nouns and geographical names are not included in the study. Some of the English-origin nouns are Canada, Toronto, CN Tower, Bloor (a street name), U-of-T ‘University of Toronto.’ Following the study on borrowing and codeswitching using the same corpus of data by Yoshizumi (2012), English nouns in the Japanese discourse of Japanese-English bilinguals are found to be borrowings. However, Park (2006) on Korean-Swedish bilingual data has reported that proper nouns behave as codeswitches, not borrowings, as revealed by morphosyntactic diagnostics, including case marking. All proper nouns are excluded from the present study.

## 5.4.2.2.3 Light Verb Construction

When either case marking or null marking is observed in ambiguous environments, those tokens are not included. The ambiguous environment observed in the present study is the light verb construction. To begin with, let us briefly describe the light verb construction in Japanese (Grimshaw and Mester 1988; Kageyama 1977; Miyagawa 1989; Miyamoto 1999; Tsujimura 1990, among others). In the Japanese light verb construction, a verbal noun is marked by the accusative case particle *o* and occurs with the verb *suru* ‘do’ as seen in the following examples in (5.16)<sup>6</sup>.

(5.16) a. *Kenta ga benkyoo o suru.*  
 (name) NOM study ACC do  
 ‘Kenta studies.’ (Literally, ‘Kenta does a study.’)

b. *Kenta ga Kanada e ryokoo o suru.*  
 (name) NOM Canada LOC travel ACC do  
 ‘Kenta travels to Canada.’ (Literally, ‘Kenta does a travel to Canada.’)

<sup>6</sup> The source of verbal nouns in Japanese can be found in native Japanese as well as foreign origins such as English, but many of the verbal nouns are Sino-Japanese, originating from Chinese (Martin 1975; Shibatani 1990).

In the following example (5.17), the verbal noun *benkyoo* is marked by the accusative case particle *o* as we have seen in (5.16a), and the preceding noun *igaku* ‘medicine’ is marked by the genitive case particle *no*, modifying the verbal noun *benkyoo*.

- (5.17) *Kenta-ga igaku-no benkyoo o suru*  
 NOM medicine-GEN study ACC do  
 ‘Kenta studies medicine.’ (Literally, ‘Kenta does a study of medicine.’)

In the Japanese light verb construction, the verbal nouns may be morphologically incorporated into *suru* (Kageyama 1977, 1991)<sup>7</sup>. According to the literature, when Sino-Japanese verbal nouns are incorporated into *suru*, the accusative case particle *o* is considered to be obligatorily deleted to form a single complex verb. As exemplified in (5.18), both verbal nouns, *benkyoo* and *ryokoo*, lack the accusative case particle *o*, compared to the examples in (5.16), and are functioning as a single complex verb with *suru*. Inflection of tense, mood, and aspect are all specified on the verb.

- (5.18) a. *Kenta ga benkyoo-suru.*  
 (name) NOM study-do  
 ‘Kenta studies.’
- b. *Kenta ga Kanada e ryokoo-suru.*  
 (name) NOM Canada LOC travel-do  
 ‘Kenta travels to Canada.’

When *benkyoo-suru* ‘to study’ functions as a single complex verb, the preceding object is marked by the accusative case particle *o* as exemplified in (5.19).

- (5.19) *Kenta ga igaku o benkyoo-suru.*  
 (name) NOM medicine ACC study-do  
 ‘Kenta studies English.’

<sup>7</sup> We have focused on Sino-Japanese verbal nouns occurring with a Japanese verb *suru*. According to Miyamoto (1999:7), Chinese verbs are borrowed as nouns and then transformed into Japanese verbs by being incorporated into *suru*.

It should be noted that the occurrence of two accusative-marked objects is not allowed in a simple clause according to the so-called Double-*O* constraint (Shibatani 1990:310). Therefore, the following example of (5.20) is not permitted.

(5.20) \**Kenta ga igaku o benkyoo o suru.*  
 (name) NOM medicine ACC study ACC do

The literature on bilingualism notes extensive cases of bilingual compound verbs (e.g. Edwards and Garner-Chloros 2007; Muysken 2000; Romaine 1986; Sankoff et al. 1990). Bilingual compound verbs are generally formed with an L2 lexical verb (e.g. English verbs) and an L1 helping verb ('pro verb'). Without exception, this is the case observed in our Japanese-English bilingual data. Two types of bilingual light verb construction occurring with the Japanese verb *suru* are observed: English noun + *suru* and English verb + *suru*. All verbal inflection is indicated by the verb *suru*. In the examples in (5.21), all English items occurring with *suru* are nouns. In (5.21a-b), the English nouns *homestay* and *audition* are marked by the accusative case particle *o*, while in (5.21c-e), the English nouns *presentation*, *discrimination*, and *discussion* are null marked. In (5.21f), the English noun *activity* is modified by the Japanese adjective *chigau* 'different' and followed by the accusative case particle *o*. When English nouns occur with the Japanese verb *suru*, they show variation in case marking. These examples are included in the study.

(5.21) English noun + *suru*

a. *homestay o suru*  
 ACC do

(CJECC/T011/50:45)

b. *audition o shi-te*  
 ACC do-CT

(CJECC/T011/296)

- c. presentation  $\emptyset$  *suru*  
ACC do  
(CJECC/T010/54)
- d. discrimination  $\emptyset$  *suru*  
ACC do  
(CJECC/T004/302)
- e. discussion  $\emptyset$  *shi-te*  
ACC do-CT  
(CJECC/O101/369)
- f. *chigau* activity  $\emptyset$  *suru* ‘to do a different activity’  
different ACC do  
(CJECC/O103/219)

In previous studies, the occurrence of English verbs followed by the Japanese verb *suru* has been reported (Azuma 1993; Nishimura 1997:98). The following examples of English verbs followed by *suru* in (5.22) are observed in our Japanese-English bilingual data. All of these English verbs are directly followed by *suru* without any case particles. In bilingual verbal constructions, some may consider that English verbs are borrowed as nominals (e.g. Moravcsik 1975). This may lead us to consider that these English verbs are behaving similarly to English nouns occurring with *suru*, assuming that they can take the accusative case particle *o*. However, none of the English verbs are case marked in our data. It would be premature to insist that English verbs are behaving like nouns without conducting a detailed analysis of our data. For example, in (5.22g-i), a direct object can occur before the English verb followed by *suru*. In (5.22g) and (5.22h), the object noun *visa* and the pronoun *sore* ‘it,’ respectively, are accusatively case marked, indicating that the English verbs, *apply-suru* and *cancel-shite*, respectively, are single complex verbs. The accusative case particle never occurs after the English verbs, in accordance with the Double-*O* constraint (Shibatani 1990) as in the above example in (5.20). In (5.22i), the Japanese noun *yosan* ‘budget’ is not case marked by the accusative case particle *o*.

This is the case of a null-marked noun in object position. The English verb *cut* followed by *suru* is functioning as a simple complex verb here. No accusative case particle is possible after the English verb *cut*.

(5.22) English verb + *suru* ‘do’

- a. stay *suru*  
do  
(CJECC/T011/320)
- b. attend *shi-te*  
do-CONJ  
(CJECC/O106/225)
- c. retire *shi-te*  
do-CONJ  
(CJECC/O106/59:12)
- d. entertain *shi-nakya*  
do-AUX  
(CJECC/O103/283)
- e. give-up *shi-cha-tte*  
do-AUX-CONJ  
(CJECC/T004/60)
- f. apply *shi-ta*  
do-PAST  
(CJECC/O106/87)
- g. visa  $\emptyset$  apply *suru*  
ACC do  
(CJECC/O106/88)
- h. *sore*  $\emptyset$  cancel *shi-te*  
it ACC do-CT  
(CJECC/O103/299)
- i. *yosan*  $\emptyset$  cut *shi-ta* *ka*  
budget ACC do-PAST Q  
(CJECC/O101/52)

Some may still claim that English verbs are borrowed as nominals and that they could take an accusative case particle. However, there is no clear linguistic justification for claiming that they are behaving like nouns in our Japanese-English bilingual data. In addition, recalling the fact that Japanese is a radical pro-drop language, ellipsis of elements in object position can also occur (Martin 1975; Shibamoto 1985). In (5.22f), it might be the case that the direct object of ‘apply-*shita* (past form of ‘apply-*suru*’) is not expressed, which could be *visa* from the context as in (5.22g). Similarly, other transitive bilingual light verb construction, as in (5.22c-e), may have a direct object noun in the previous context, but it is unexpressed. For these reasons, tokens of an English verb followed by the Japanese verb *suru* are not included in the present study.

Although the examples of English nouns and English verbs occurring with *suru* have been presented, we have observed examples of English items that can be either nouns or verbs, such as *dance*, *volunteer*, *show*, and so on, as exemplified in (5.23-26). In (5.23a), *dance* is directly followed by *shi-tari* ‘do-REP’ and it is ambiguous with regard to whether it functions as a single complex verb with *suru* or if the accusative case particle *o* is omitted (i.e. null-marked noun). In (5.23b), the accusatively case marked *dance* is modified by its preceding genitive case marked *local*. Therefore, *dance* in (5.23b) is a noun and should be included in the study. However, the category of *dance* in (5.23a) is ambiguous, leading us to exclude it from the study. Similarly, the category of *volunteer* in (5.24a-b) is not clear from the context whereas that of *volunteer* in (5.24c) is considered as a null-marked noun as it is modified by *kodomo-no* ‘child-GEN’. Therefore, the examples in (5.24a-b) are excluded from the study while the example in (5.24c) is included. Both examples in (5.25) are included as *show* is a noun functioning as a direct object, which refers to a *kimono* (traditional Japanese clothing) fashion show. Inclusion of these contexts is justified on the grounds that *show* in (5.25a) is recoverable from the ambient discourse context,

while the null marked *show* in (5.25b) appears with a modifier (*souitta* ‘that kind of’). Lastly, in the case of *trouble* in (5.26a), it is not clear if it is a null-marked object noun or part of a single complex verb formed with *suru*. This example is not included. The English word *manicure* in (5.26b) can be a noun or a verb in English. In this example, however, it is a verb occurring with *suru* since its object *tume* ‘nail’ also appears. Therefore, *manicure-suru* is a single complex verb and the accusative particle cannot occur after *manicure*. The example is not included in the study.

(5.23) *dance*

a. dance *shi-tari*  
do-REP (CJECC/O103/340)

b. local-*no* dance  $\emptyset$  *suru*  
GEN ACC do (CJECC/O103/345)

(5.24) *volunteer*

a. volunteer *shi-te-kudasar-u*  
do-CT-give(HON)-PRES (CJECC/T011/261)

b. volunteer *shi-te-ki-ta*  
do-CT-come-PAST (CJECC/T018/316)

c. *kodomo-no* volunteer *shi-te-ki-ta*  
child-GEN do-CT-come-PAST (CJECC/T018/312)

(5.25) *show*

a. show  $\emptyset$  *shi-ta-n-desu*  
ACC do-PAST-SE-CPL (CJECC/T011/367)

b. *souitta* show *shi-ta-n-desu*  
that.kind.of do-PAST-NM-CPL (CJECC/T011/370)

(5.26) Other verbs

a. trouble *shi-te*  
do-COT

(CJECC/T004/253)

b. *tume*  $\emptyset$  *makka ni* *manicure-shi-te*  
nail ACC red LOC do-CT  
‘[They] manicure [their] nails in red and...’

(CJECC/O101/317)

In sum, when English items occurring with *suru* are observed, the tokens included in the present study are only those that are unambiguously nouns, regardless of whether or not they are case marked or null marked. All tokens that are not nouns or that are ambiguous with regard to their categorial status (i.e. noun or verb) are not included.

The same criteria are used when Japanese tokens are extracted. Matsuda (1995), who investigated variation on the accusative *o* case marking in Tokyo Japanese, also mentions that it is controversial to claim that any verbal noun directly followed by the verb *suru* is lacking the accusative case particle *o*. Although it may be the case that (Sino-Japanese) verbal nouns show variable accusative case *o*-marking before the verb *suru*, it may also be the case that so-called S(Sino-Japanese)-incorporation (Kageyama 1982) has taken place by applying a rule of obligatory “*o*-deletion,” as in the earlier examples of (5.18). As Matsuda (1995) mentions, this point requires further analysis before including these ambiguous cases in our study. Therefore, following Matsuda (1995), and based on the bilingual examples observed in our data presented above, the study includes the Japanese light verb construction only when there is a linguistically defensible reason for claiming that the object items in question are unequivocal nouns that can show variable case-marking forms.

## 5.5 Hypothesis

As mentioned in §1.5.3, the present study compares four types of nouns: English-origin nouns and Japanese nouns in heritage Japanese, and Japanese nouns and loanwords in homeland benchmark Japanese. Given the absence of any analogous system of case marking in English, we hypothesize that sustained contact with English in Canada may result in the attenuation of variable case marking in Japanese-English bilingual speech (Clyne 2003; Huffines 1989, 1991; Larmouth 1974; Meyerstein 1969; Preston 1986; Polinsky 1995, 1997). We pursue this inference by focusing on the variable case marking of English-origin nouns in the discourse of Japanese-English bilinguals in Canada in order to ascertain whether the case marking system used by this group of speakers diverges significantly from that used by speakers of the monolingual homeland variety. Furthermore, we will investigate whether or not English-origin nouns in otherwise Japanese discourse are behaving similarly to or differently from Japanese nouns in unmixed Japanese discourse and then compare the patterning of Japanese nouns and loanwords in homeland benchmark Japanese to assess contact-induced change. An important working principle underpinning the ensuing analysis is that the status of variable case marking in the heritage and homeland varieties of Japanese can only be determined by systematically examining multiple comparison points.

## 5.6 Factor Groups

After extracting all the relevant tokens, each of them is coded according to a number of social and linguistic factors that have been reported in the literature to constrain variant choice (i.e. case marking or null marking). All coded data is analyzed by means of GoldVarb X (Sankoff et al. 2005). The logistic regression procedure in this program reveals which

independent variables make a significant contribution to variant choice when all of them are considered simultaneously. Social factors include length of stay in Canada for heritage Japanese speakers, age for monolingual Japanese speakers of the homeland benchmark variety, and sex for both cohorts. Linguistic factor groups to be examined are verbal adjacency, presence of focus particle, presence of sentence-final particles, word order, and clause type. Each of these factors represents a testable hypothesis. These factors are described below.

### 5.6.1 Social Factors

As discussed in the previous chapter on the variable realization of subject pronouns, the speaker's length of stay in Canada was examined for heritage Japanese to test the hypothesis that the longer speakers are exposed to the majority language, English, in a contact setting, the more likely they are to exhibit structural change in their speech patterns as a result of lengthy contact with English (Thomason 2001; Thomason and Kaufman 1988; Winford 2003). We examined speakers' age in the homeland benchmark variety, which, by virtue of the inclusion of an apparent-time component, enables us to ascertain whether there is a possible change in progress in this variety in relation to the variable case marking of subject nouns and direct object nouns. Speakers' sex was also included to see if similar findings to those reported in previous studies could be observed in the data: Shibamoto (1985) reports that null marking occurs more frequently among female speakers than male speakers. It has also been reported that male speakers retain particles even more with direct object nominals than subject nouns (Shibamoto 1985). Matsuda (1995), discussing the variable accusative marker *o*, shows that omission of the accusative case marker *o* is significantly favoured by female speakers compared to males.

## 5.6.2 Linguistic Factors

### 5.6.2.1 Verbal Adjacency

According to Tsusui (1984:132), “[un]less the speech is very formal, the ellipsis of the *o* of an NP-*o* in a sentence is natural if the NP-*o* is immediately followed by the predicate of the sentence.” Saito (1983) has also made the same claim regarding the ellipsis of the accusative case particle *o*. In keeping with this claim, several quantitative studies have reported that ellipsis of the accusative case particle *o* is favoured when direct object NPs are immediately followed by the predicate (Fujii and Ono 2000; Matsuda 1995). Fry (2003) examined the effect of verb adjacency on not only direct object nouns, but also subject nouns with respect to ellipsis of case markers. For object nouns, Fry’s results support the previous findings for case marking including not only the accusative case marker *o* but also others such as the nominative case marker *ga* for nominative objects and the topic marker *wa*, which are included in the present study as well. However, Fry reports that there is no statistically significant difference for the ellipsis of the case marker of subject nouns in relation to verbal adjacency.

Considering the previous findings for the ellipsis of the accusative case particle *o* in the literature, we hypothesize that case-particle ellipsis may be promoted when the subject noun is directly adjacent to the predicate of the sentence as in (5.27a-b). The nouns *kodomo-tachi* ‘children’ and *o-okme* ‘rice’ are directly adjacent to the predicate. Conversely, the adverb *sugoi* ‘very’ is inserted between the subject noun *kiri* ‘fog’ and the verb *kokunatyau* ‘to get thickened’ in (5.28a) while the adverb *betsuni* ‘particularly’ appears between the direct object noun *tai* ‘seabream’ and the verb *kuu* ‘to eat’ in (5.28b).

(5.27) Adjacent

- a. *Kodomo-tachi*  $\emptyset$  *yu-tte-ta.*  
 child-PL NOM say-ASP-PAST  
 ‘[My] children were saying [so].’

(CJECC/T018/206)

- b. *O-kome*  $\emptyset$  *ka-tta* *ne.*  
 HON-rice ACC buy-PAST SFP  
 ‘[I] bought rice.’

(CJECC/O106/211)

(5.28) Non-adjacent

- a. *Kiri* *ga* *sugoi* *ko-ku-natyau* *none.*  
 fog NOM very thicken-CT-AUX SFP  
 ‘The fog becomes much thicker.’

(CJECC/O101/11)

- b. *Tai* *o* *betsuni* *ku-u-wake-ja-na-i-n-da* *kedo.*  
 seabream ACC particularly eat-PRES-REASON-COP-NEG-PRES-SE-COP but  
 ‘[We] do not eat seabream in particular, though.’

Table 5.1: Hypothesis for the effect of verbal adjacency on variant choice

Hypothesis	Favours null marking (on direct object)	Reference
Adjacent	✓	Fujii and Ono (2000); Fry (2003); Matsuda (1995); Saito (1983); Tsutsui (1984)

### 5.6.2.2 Presence of Focus Particle

Japanese nouns can occur with a particle that functions as a focus device. The focus particle is preceded by the noun and followed by other particles such as the nominative case particle *ga*, the accusative case particle *o*, the topic marker *wa*, and so on. The nouns followed by focus particles receive meanings such as *only*, *among others*, *even*, etc. With regard to case-marking ellipsis, Matsuda (1995) has found that the ellipsis of the accusative case particle *o* is strongly associated with the presence of the focus particle co-occurring with the same object noun. Each token was coded according to the presence or absence of focus particles. The following examples in (5.29a-b) demonstrate the presence of a focus particle following English

nouns. In (5.29a), the subject noun phrase *ensouryou shuunyuu* ‘income of performance fee’ is followed by the focus particle *toka* ‘among others’ and the nominative case particle *ga* is omitted, whereas in (5.29b) the direct object noun *mitsumori* ‘estimate’ is followed by the focus particle *bakari* ‘only’ and the accusative case particle *o* is omitted.

(5.29) Presence

a. *Ensou-ryou shuunyuu toka Ø kekkoo tama-tari-shi-te ne.*  
 performance-fee incomes FP NOM fairly pile.up-REP-DO-CT SFP  
 ‘Income from performance fees piles up surprisingly.’  
 (CJECC/O122/259)

b. *Mitsumori bakkari Ø tor-e-ru.*  
 estimate FP(only) ACC get-POT-PRES  
 ‘[We] can only get an estimate.’  
 (CJECC/O110/156)

Table 5.2: Hypothesis for the effect of presence of focus particle on variant choice

Hypothesis	Favours null marking	Reference
Presence	√	Matsuda (1995)

### 5.6.2.3 Presence of Sentence-final Particle

In Japanese, sentence-final particles occur in a sentence-final position and interjectional particles can occur freely within a clause, including a verb-final position (Shibatani 1990:334). By adding these particles in a verb-final position, the sentence acquires specific pragmatic functions. For example, the particle *yo* is used to call the attention of the listener to the information, as in (5.30a), while the particle *ne* can be used to express the speaker’s agreement, as in (5.30b).

- (5.30) a. *Sorede kouyuu hanasshi o kiki-mashita-desu yo.*  
 and.then this story ACC listen-PAST-COP SFP  
 ‘And then I listened to this story.’

(CJECC/004/477)

- b. *ningen-no nanka, kou, chiisasa o kanji-masu ne.*  
 human.being-GEN somehow INT smallness ACC feel-PRES SFP  
 ‘I somehow feel the smallness of human beings.’

(CJECC/001/188)

According to Masunaga (1988), the omission of case-marking particles is correlated with the presence of a final particle. Attachment of a final particle makes a verb in the same sentence more focused, which makes the other elements defocused. As a result, ellipsis of the case-marking particle of the defocused element can take place, as shown in the following examples in (5.31-32) presented in Masunaga (1988:148). In both sentences with a final particle *-zo* or *-yo*, the subject nouns *onnanoko* ‘boy’ and *otokonoko* ‘girl’ can drop the nominative case *ga*.

- (5.31) *Onnanoko Ø kita zo*  
 girl came yo  
 ‘A girl came.’

(source: Masunaga 1988:148)

- (5.32) *Brondo no otokonoko Ø Taro o nagutta zo*  
 blond GEN boy ACC hit yo  
 ‘A blond boy hit Taro.’

(source: Masunaga 1988:148)

Following this assumption, Matsuda (1995) conducted a quantitative analysis of the correlation of accusative case marking with the presence of final particles found in his data. He found that the accusative case particle *o* is more often dropped when there is a final particle in verb-final position. Matsuda mentions that his data duplicate Masunaga’s findings “in that [his] analysis detected a meaningful correlation between the presence of FP and the likelihood of

absence of *o*” (p.174). As final particles are observed to be used widely in our data, we will also examine the effect of sentence-final particles on the variability in case marking of subject and object nouns. Following previous research, it is hypothesized that presence of a final particle favours case-marking omission. All tokens are coded according to the presence or absence of final particles in our data.

In (5.33a), the sentence-final particle *none* is attached in verb-final position and in the same sentence the subject noun *danna-san* ‘husband’ is followed by the nominative case particle *ga*. Similarly, in (5.33b), the final particle *yone* is attached in verb-final position and the subject noun *shuugiin-senkyo* ‘election for the House of Representatives’ is null marked. Conversely, the final particle *ne* occurs in verb-final position, as in (5.33c), but the direct object nouns *osushi* is marked by the accusative case marker *o*. In (5.33d), the direct object noun *o-tanjoubi-kai* ‘birthday party’ is null marked. Finally, in (5.33e), the verb *hajimatta* ‘started’ is followed by a tag phrase *-desho*. It is not a particle, but a tag expression soliciting the listener’s attention and functioning similarly to sentence-final particles. Accordingly, tokens such as (5.33e) are also included.

(5.33) Present

- a. *Danna-san ga waru-katta none.*  
 husband-Mr. NOM bad-PAST SFP  
 ‘[Her] husband was bad.’

(CJECC/T010/148)

- b. *Shuugiin-senkyo ∅ hajimari-masu yone.*  
 house.of.representatives-election NOM start-DECL SFP  
 ‘The election for the House of Representatives will start.’

(CJECC/T004/257)

c. *O-sushi o tsuku-ru-n-da ne.*  
HON-sushi ACC make-PRES-SE-COP SFP  
 ‘[We] make sushi.’

(CJECC/T018/146)

d. *O-tanjoubi-kai ∅ san.kai shi-ta no.*  
HON-birthday-party ACC three.times do-PAST SFP  
 ‘[We] did a birthday party three times.’

(CJECC/O103/252)

e. *Sorede zaigai-senkyo ga hajima-tta desho.*  
consequently overseas-election NOM start-PAST TAG  
 ‘That’s why the overseas election started, didn’t it.’

(CJECC/T004/254)

Table 5.3: Hypothesis for the effect of presence of sentence-final particle on variant choice

Hypothesis	Subject		Direct Object	
	Favours null marking	Reference	Favours null marking	Reference
<b>Presence</b>	✓	Masunaga (1988)	✓	Masunaga (1988); Matsuda (1995)

#### 5.6.2.4 Word Order

It has been suggested that omission of a case particle is caused by the scrambling of constituents in the sentence (Tamori 1977). In Takano’s (1998) study conducted within a variationist framework, it was found that particle ellipsis of *wa/ga* is favoured when subject nouns are dislocated and occur in post-predicate position. Prior to Takano’s study, Shibatani (1990) had assumed that *ga* tends to be omitted in the canonical word order of Japanese, where the subject NP appears before the predicate<sup>8</sup>. As for object nouns, Fujii and Ono (2000) have observed that case-marking ellipsis is more likely to occur when the object nouns are in pre-predicate position, whereas accusative case *o*-marking is more frequent in post-predicate position. In our data, all tokens are coded according to whether the nouns are in pre-predicate position

<sup>8</sup>Shibatani (1990) also found that this tendency is inhibitive in women’s particle ellipsis.

(SV/OV) as in (5.34) or in post-predicate position (VS/VO) as in (5.35a-b). In example (5.35a), the subject noun of the predicate *hayai desu* ‘to be quick’ is *hannou* ‘reaction,’ but it is dislocated into a post-predicate position. In example (5.35b), the direct object of the verb phrase *tsurete-kite-kudasai* ‘Please bring’ is *otomodachi* ‘friend,’ which is dislocated into a post-predicate position.

(5.34) Pre-predicate

a. *Atashi wa ito to hari Ø mo-tte-na-i.*  
 I TOP string and needle ACC have-CT-NEG-PRES.  
 ‘I don’t have string and needles.’

(CJECC/O121/148)

(5.35) Post-predicate

a. *Hayai desu yone, hannou ga.*  
 quick COP(PRES) SFP reaction ACC  
 ‘Reaction is quick.’

(CJECC/T011/139)

b. *Tsure-te-ki-te kudasai, o-tomodachi o ne.*  
 take-CT-come-CT please.give HON-friend ACC INT  
 ‘Please bring your friends.’

(CJECC/T011/272)

Table 5.4: Hypothesis for the effect of word order on variant choice

Hypothesis	Subject		Direct Object		
	Favours null marking	Reference	Favours overt marking	Favours null marking	Reference
<b>Pre-predicate (SV/OV)</b>	✓	Shibatani (1990)		✓	Fujii and Ono (2000)
<b>Post-predicate (VS/VO)</b>	✓	Tamori (1977); Takano (1998)	✓		Fujii and Ono (2000)

## 5.6.2.5 Clause Type

Matsuda (1995) examined the effect of the clause type ('degree of embeddedness') on variable accusative case *o*-marking for object nouns in order to assess language change in Tokyo Japanese. Following Hooper and Thompson's (1973, cited in Matsuda 1995) finding that "in syntax, a smaller number of transformational rules are applicable to embedded clauses than to main clauses" (p. 96), Matsuda tested to see whether or not the same tendency could be observed for accusative case marking. He found that *o*-ellipsis is disfavoured in embedded clauses when compared to main clauses and semi-embedded clauses. To confirm the effect of clause type in the present data, each token was coded as either a main clause, (5.36), or as a subordinate clause (e.g. adverbial clause, relative clause, etc.), as in (5.37a-c).

## (5.36) Main Clause

*Yappa ii-tai-koto o i-u.*  
 after.all say-DES-fact ACC say-PRES  
 'After all, [I] say what I want to say.'

(CJECC/O102/309)

## (5.37) Subordinate Clause

## a. Adverbial Clause

*Soto-ni teki* *Ø tsuku-reba, ichiban matoma-ru-wake-desu yone.*  
 outside-LOC enemy ACC make-COND the.first.place be.unified-PRES-REASON-COP SFP  
 'If [they] make enemies outside, [they] get unified the most.'

(CJECC/T004/114)

## b. Complement

*De, ohana* *Ø shi-you to omo-tte.*  
 then flower.arranging ACC do-VOL COMP think-CT  
 'Then, [I] thought that I would do flower arranging.'

(CJECC/123/226)

## c. Relative Clause

[*Warui koto* *Ø shi-teru*] *hito ga takusan ite-ru yan.*  
 bad thing ACC do-PROG person NOM a.lot existence-PRES TAG  
 'There are a lot of people who are doing bad things.'

(CJECC/O112/110)

Table 5.5: Hypothesis for the effect of clause type on variant choice

Hypothesis	Favours null marking	Reference
Embedded	✓	Matsuda (1995)

## 5.7 Results

### 5.7.1 Overall Distribution

Up to 25 tokens of English-origin nouns and Japanese nouns were extracted in subject position and direct object position in heritage Japanese. In the same way, up to 25 tokens of Japanese nouns and loanwords were extracted in subject position and direct object position respectively from the homeland benchmark variety. Loanwords include only English-origin nouns<sup>9</sup>. A total of 1132 tokens were extracted for subject nouns, whereas a total of 1183 tokens were extracted for direct object nouns.

Table 5.6 presents the overall rates of variable subject marking. With all types of subject nouns, the most frequent marker is the nominative case marker *ga*. With English-origin nouns in heritage Japanese, and Japanese nouns and loanwords in the homeland benchmark variety, the second most frequent occurrence is null marking, followed by the topic marker *wa*. NRLI (1955:113-114) has reported the same order of occurrence in their conversational data, although the proportion is different: nominative case marker *ga* (33.5%), null marking (28.9%), and topic

<sup>9</sup> Although there are lots of loanwords in Japanese from other Western languages such as Portuguese, Dutch, and German, the present study focuses on English-origin nouns only for a comparison with English-origin nouns in heritage Japanese. However, in the course of extraction, 12 Western language-origin loanwords were found in the direct object position as follow:

Portuguese:	<i>rinpa</i> ‘lymph’	<i>pan</i> ‘bread’ (from <i>pan</i> )	<i>tabako</i> ‘tobacco’ (from <i>tabaco</i> )
Dutch:	<i>randoseru</i> ‘a school backpack’ (from <i>ransel</i> )	<i>panki</i> ‘paint’ (from <i>penki</i> )	
	<i>biiru</i> ‘beer’ (from <i>bier</i> )	<i>kan</i> ‘can’ (from <i>kan</i> )	
German:	<i>baito/arubaito</i> ‘part-time job’ (from <i>Arbeit</i> )		

marker *wa* (23.7%)<sup>10</sup>. In contrast with English nouns in heritage Japanese, Japanese nouns in the heritage variety exhibit a slightly different order, with a reversal of the rates for null marking and the topic marker *wa*.

Table 5.6: Overall rates of nouns for variable case marking in subject position

Subject	Heritage Japanese				Homeland Japanese			
	English-origin nouns		Japanese nouns		Japanese nouns		Loanwords	
particle	%	N	%	N	%	N	%	N
<b><i>ga</i></b>	<b>63.5</b>	165	<b>61.1</b>	231	<b>60.0</b>	233	<b>48.5</b>	48
<b><i>wa</i></b>	<b>13.8</b>	36	<b>20.9</b>	79	<b>15.9</b>	63	<b>15.2</b>	15
<b>Ellipsis</b>	<b>22.7</b>	59	<b>18.0</b>	68	<b>25.1</b>	99	<b>36.4</b>	36
Total	100	260	100	378	101	395	100.1	99

In our data, the nominative case marker *ga* is used to indicate the subject nouns in English-origin nouns at 63.5% and Japanese nouns in heritage Japanese at 61.1%. Japanese nouns in homeland Japanese are marked by *ga* at 60%, while it is used least frequently in loanwords in homeland Japanese at 48.5%. Null marking occurs most frequently with loanwords at 36.4% in homeland Japanese, whereas Japanese nouns in heritage Japanese have the lowest rate of null marking at 18%. A comparison of null marking of loanwords in homeland Japanese (36.4%) and English-origin nouns (22.7%) in heritage Japanese shows that the rate of null marking is higher in the homeland variety. Taken together, these preliminary rate comparisons suggest that contact with English in the heritage variety has not resulted in the attenuation or erosion of the case-marking system for subject nouns. Had rate differences been indicative of the loss of overt case marking, implicating potential contact effects, higher rates of particle ellipsis would be expected to occur in heritage Japanese, which is not confirmed by the data depicted in Table 5.6.

<sup>10</sup> The remaining 15.9% is accounted for by other particles such as *mo*, *tte datte*, and *demo*

Table 5.7: Overall rates of nouns for variable marking in direct object position

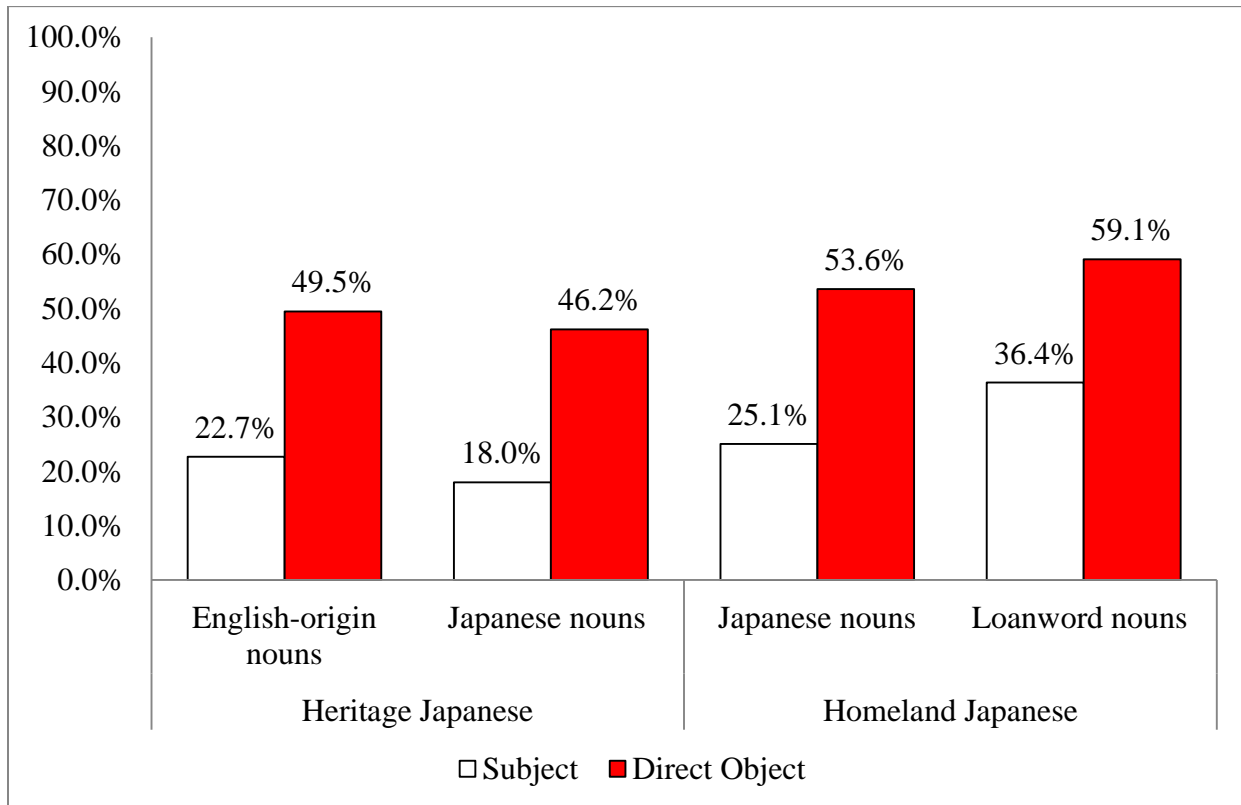
Direct Object	Heritage Japanese				Homeland Japanese			
	English-origin nouns		Japanese nouns		Japanese nouns		Loanwords	
Particle	%	N	%	N	%	N	%	N
<i>o</i>	<b>37.9</b>	114	<b>29.4</b>	112	<b>26.4</b>	102	<b>22.6</b>	26
<i>ga</i>	<b>7.6</b>	23	<b>13.9</b>	53	<b>8.8</b>	34	<b>10.4</b>	12
<i>wa</i>	<b>5.0</b>	15	<b>10.5</b>	40	<b>11.1</b>	43	<b>7.8</b>	9
<b>Ellipsis</b>	<b>49.5</b>	149	<b>46.2</b>	176	<b>53.6</b>	207	<b>59.1</b>	68
Total	100	301	100	381	99.9	386	99.9	115

Table 5.7 displays the overall rates of direct object marking. In contrast with subject nouns, where the most frequent particle is the nominative case marker *ga*, null marking predominates in the case of direct object nouns for all types of nouns. Rates of null marking are below 50% in heritage Japanese (English-origin nouns = 49.5% and Japanese nouns = 46.2%) while it occurs more frequently in the homeland benchmark variety (Japanese nouns at 53.6% and loanwords at 59.1%). Again, these preliminary proportional comparisons do not militate in favour of the erosion of the case-marking system for object nouns in the heritage variety. For overt marking on direct object nouns, the second most frequent occurrence is the use of the accusative case marker *o* in all types of nouns, followed by either the nominative case marker *ga* used in the nominative object construction or the topic marker *wa*. These results confirm those reported in previous studies targeting informal colloquial speech (Martin 1975; Murasugi and Sugisaki 2008; Shibatani 1990; Tsutsui 1984; Matsuda 1995, and others) in that post-nominal case particles can be variably omitted.

Figure 5.1 demonstrates the rate of null marking on nouns by position: subject nouns and direct object nouns. First, Japanese nouns in homeland benchmark Japanese trigger null marking 25.1% of the time in subject position, while rates of null marking are higher in direct object

position at 53.6%. Loanword nouns show null marking 36.4% of the time in subject position while they are also null marked more frequently in direct object position at 59.1%. On the other hand, in heritage Japanese, English-origin nouns occur without a case marker in subject position 22.7% of the time whereas they occur as null more frequently in direct object position, 49.5% of the time. Japanese nouns in heritage Japanese show null marking in subject position 18.0% of the time while they are null marked more frequently in direct object position at 46.2%. Regarding the difference in variable case marking of subjects and direct objects between comparison varieties of Japanese, a Chi-square test reveals a statistically significant distinction (i.e.  $p=0.001022$  in subject position and  $p=0.048986$  in direct object position).

Figure 5.1: Overall rates of null marking on nouns in subject position and direct object position



According to the literature (Martin 1975:50), the object case marker *o* drops more frequently than the subject case marker *ga*. Fry (2003) discusses nouns in direct object position which apparently show a higher rate: 53% for direct object nouns and 33% for subject nouns. Although the rates of null marking on nouns observed in our data do not exactly correspond to those found by Fry, our results share the same direction of effect, with object nouns exhibiting a greater preference for the null variant than subject nouns<sup>11</sup>.

Although the differences in rates of occurrence are found to be statistically significant in the present study, this difference is not necessarily the product of language contact. As discussed in the literature (Poplack and Tagliamonte 2001; Torres-Cacoullos and Travis 2010), rates can fluctuate for all kinds of non-linguistic reasons. In other words, rate differentials are not reliable diagnostics of underlying structural differences between varieties. In order to detect such differences, inspection of the grammar underlying variable case-marking patterns needs to be effected. This will enable us to assess the extent to which underlying quantitative patterns of variable case marking are isomorphic or divergent across the two comparison varieties of Japanese. In the following section, we subject the data to a multivariate analysis.

### 5.7.2 Multivariate Analysis

In this section, we present the results obtained from a multivariate analysis of the data. This analysis is carried out in order to assess the linguistic and social factors which are significant in

---

<sup>11</sup> In Fry (2003), particle ellipsis is described as the context where any particle drops from the head noun with a specific argument role. Fry's (2003) study is different from the present study in that we focus on variability in case marking, or its ellipsis with nouns in subject position and object position, and we do not include any particles that do not function as case markers. Fry's (2003) data includes particles that are used as postpositions, such as *de* 'at/in' and *kara* 'from,' which correspond to prepositions in English, as long as they appear in the subject/agent argument role position and direct object argument role position. Particles such as *kara* and *de* can never be omitted. Differences between the results reported in the present study and those reported by Fry (2003) are evidently due to different data configurations.

determining variant choice. The comparison of constraint hierarchies generated by the logistic regression analyses will enable evaluation of whether the same underlying grammar is shared by the two varieties of Japanese with respect to the two targeted variables: variable case marking of subject nouns and variable case marking of direct object nouns.

### 5.7.2.1 Variable Case Marking on Nouns in Subject Position

Table 5.8 presents the overall results of the multivariate analysis of social factors of null marking on subject nouns: English nouns and Japanese nouns in heritage Japanese, and Japanese nouns and loanwords in homeland Japanese. As indicated by the corrected mean, the overall rate of null marking on subject nouns is higher in homeland Japanese (.251 for Japanese nouns and .364 for loanwords) than in the heritage variety (.180 for Japanese nouns and .227 for English-origin nouns). In the homeland benchmark variety, null marking in Japanese nouns is favoured by younger speakers with a probability of .60. This is an intriguing apparent-time difference, potentially pointing to a possible change in progress. Because no recent study has replicated this finding, it seems expedient to reserve judgement for the moment on whether this effect constitutes a bona fide change in progress.

Table 5.8: Multivariate analysis of the contribution of social factors selected as significant to the probability of null marking on subject nouns; factor groups not selected as significant are in square brackets.

	HERITAGE JAPANESE						HOMELAND JAPANESE					
Subject	English-origin nouns			Japanese nouns			Japanese nouns			Loanwords		
Corrected mean	.227			.180			.251			.364		
Token N	59/260			68/378			99/395			36/99		
	FW	%	Token N	FW	%	Token N	FW	%	Token N	FW	%	Token N
<b>LENGTH OF STAY IN CANADA</b>												
Longer	.58	28.1	36/128	[.56]	21.6	43/199	N/A			N/A		
Shorter	.43	17.4	23/132	[.43]	14.0	25/179	N/A			N/A		
Range	15											
<b>AGE</b>												
Younger	N/A			N/A			.60	32.3	64/198	[.49]	36.2	21/58
Older	N/A			N/A			.40	17.8	35/197	[.52]	36.6	15/41
Range	20											
<b>SEX</b>												
Female	[.51]	23.1	31/134	[.51]	18.4	35/190	[.53]	25.5	50/196	[.42]	29.5	13/44
Male	[.49]	22.2	28/126	[.49]	17.6	33/188	[.47]	24.6	49/199	[.56]	41.8	23/55
Range												

Regarding the social conditioning of variant choice in heritage Japanese, null marking on English-origin nouns is favoured by heritage Japanese speakers who have been resident in Canada for a longer period, with a probability of .58. Length of stay in Canada is not selected as significant for Japanese nouns in heritage Japanese, although speakers who have been in Canada for a longer period are more likely to use null marking, as indicated by the direction of the effect for this factor group. Since it has been claimed that structural change is promoted by a long period of contact between two languages (Thomason 2001; Thomason and Kaufman 1988; Winford 2003), this finding may suggest that due to the influence of English, the speakers who

have been in Canada for a longer period tend to drop subject markers from nouns, in particular English-origin nouns, compared to bilingual speakers with a shorter period of residence.

Speaker sex is not selected as significant in all types of nouns in heritage and homeland Japanese, in spite of claims in the literature that null marking is more likely to be used by female speakers.

Table 5.9 synthesizes the results of the multivariate analysis of linguistic factors for nouns in subject position<sup>12</sup>. Inspection of the results for Japanese nouns in the homeland variety shows that the presence of a focus particle contributes the strongest effect to variant choice, as assessed by the range value of 45. Japanese nouns are preferentially null marked when they are followed by a focus particle, with a probability of .91, as reported in Matsuda (1995) for his data based on Tokyo Japanese<sup>13</sup>. The second most significant factor for Japanese nouns in homeland Japanese is verbal adjacency with a range value of 15, followed by sentence-final particle with a range value of 12. Japanese nouns are null marked when they are directly followed by their predicate with a probability of .54 and when the sentence-final particle is present with a probability of .56. These findings are consistent with previous studies (Fry 2003; Fujii and Ono 2000; Matsuda 1995; Masunaga 1988; Saito 1983; Tsutsui 1984). As for loanwords in homeland Japanese, the focus particle factor group is also selected as significant with a range value of 28. Presence of a focus particle favours null marking with a probability of .74. Verbal adjacency and sentence-final particle are not selected as significant, and the direction of effect for verbal adjacency is reversed for homeland loanwords in comparison with the effect for Japanese nouns. Clause type

---

<sup>12</sup> Word order was not included in the multivariate analysis because non-canonical orders such as Verb-Subject are rarely observed. As occurrences of verb-subject and verb-direct object are too sparse in our data to investigate the effect of this factor, word order was removed from the multivariate analysis (see Guy 1988 for the details of rates of over 95% or below 5%).

<sup>13</sup> However, recall that Matsuda (1995) investigated variable accusative case marking.

was not selected as significant for either type of noun in homeland Japanese; however, the direction of the effect is identical.

Table 5.9: Multivariate analysis of the contribution of linguistic factors selected as significant to the probability of null marking on subject nouns; factor groups not selected as significant are in square brackets.

	HERITAGE JAPANESE						HOMELAND JAPANESE					
Subject	English-origin nouns			Japanese nouns			Japanese nouns			Loanwords		
Corrected mean	.227			.180			.251			0.364		
Token N	59/260			68/378			99/395			36/99		
	FW	%	Token N	FW	%	Token N	FW	%	Token N	FW	%	Token N
<b>FOCUS PARTICLE</b>												
Presence	.91	73.3	11/15	.93	70.0	7/10	.91	74.1	20/27	.74	61.5	8/13
Absence	.47	19.6	48/245	.48	16.6	61/368	.46	21.5	79/368	.46	32.6	28/86
Range	<b>44</b>			<b>45</b>			<b>45</b>			<b>28</b>		
<b>VERBAL ADJACENCY</b>												
Adjacent	[.52]	24.7	43/174	.60	23.4	52/222	.54	28.0	78/281	[.49]	36.6	26/71
Non-adjacent	[.46]	18.9	14/74	.35	10.3	15/146	.39	20.2	21/104	[.53]	38.5	10/26
Range				<b>25</b>			<b>15</b>					
<b>SENTENCE-FINAL PARTICLE</b>												
Presence	[.52]	25.7	28/109	[.58]	21.7	38/175	.56	28.6	60/210	[.52]	38.9	21/54
Absence	[.49]	20.7	31/150	[.43]	14.4	28/195	.44	20.7	38/184	[.47]	34.1	15/44
Range							<b>12</b>					
<b>CLAUSE TYPE</b>												
Main	[.51]	24.4	33/135	[.51]	18.9	40/212	[.51]	27.6	67/243	[.53]	37.7	23/61
Subordinate	[.49]	21.0	26/124	[.48]	17.1	28/164	[.49]	20.7	31/150	[.46]	34.2	13/18
Range												

On the other hand, inspection of the results in heritage Japanese shows that the presence of a focus particle also contributes the strongest effect to variant choice for both Japanese nouns and English-origin nouns, as assessed by the range value of 45 and 44, respectively. Null marking on subject nouns is favoured when a focus particle appears with a probability of .93 for Japanese nouns and .91 for English-origin nouns. For Japanese nouns, verbal adjacency is also

selected as significant with a range value of 25. When Japanese nouns are directly followed by their predicate, null marking occurs with a probability of .60. This factor is not selected as significant for English-origin nouns, but the same direction of the effect obtains. The sentence-final particle factor group is not selected as significant for English-origin nouns and Japanese nouns in heritage Japanese, but the direction of the effect is identical. Finally, clause type was not selected as significant for any type of noun in heritage Japanese, but the identical direction of the effect is observed.

Summarizing, the focus particle factor group is the key determinant of variant choice for all types of nouns. More importantly, the constraint hierarchies that each factor is comprised of are parallel in all factor groups across all types of nouns between heritage Japanese and the homeland benchmark variety. No differences in terms of the direction of effects are observed between the two varieties, with the exception of a mitigated and reversed direction of effect in verbal adjacency for loanwords in the homeland variety. Comparison of constraint hierarchies suggests that, in general, the same underlying grammars are operative in heritage and homeland Japanese.

However, recall that length of stay in Canada was selected as significant for English-origin nouns in heritage Japanese, while age was selected as significant for Japanese nouns in homeland Japanese. From these findings, we now inquire whether or not heritage Japanese speakers show a different underlying grammar depending on the length of stay in Canada and whether or not there is an ongoing language change with regard to variable case marking on subject nouns in the homeland benchmark. If there are significant differences between speakers who have been in Canada for a longer period and those who have been in Canada for a shorter period, and furthermore, if English-origin nouns according to the speaker's length of stay in

Canada are behaving differently from Japanese nouns in the homeland benchmark, then it may be surmised that there is some possible English influence. In the homeland variety, on the other hand, if younger speakers and older speakers show differences in variable case marking on Japanese subject nouns, it may be considered as a possible language change.

First, Table 5.10 presents the results of a multivariate analysis of null marking on English-origin subject nouns for heritage Japanese speakers according to the length of stay in Canada. Only the focus particle factor group is selected as significant with a range value of 34 for the heritage speakers who have stayed in Canada for a longer period and with a range value of 50 for the speakers who have stayed in Canada for a shorter period. Null marking is favoured in the presence of a focus particle in both cohorts. Although none of the remaining factor groups are selected as significant, the same direction of effect and neutralized effects are observed, suggesting no significant difference. The comparative analysis according to the length of stay in Canada in heritage Japanese indicates that a similar underlying grammar is shared by both cohorts for variable case marking on English-origin subject nouns, with a strong effect associated with the focus particle factor group.

Table 5.10: Multivariate analysis of the contribution of linguistic factors selected as significant to the probability of null marking on English-origin subject nouns according to length of stay in Canada for heritage Japanese speakers; factor groups not selected as significant are in square brackets.

	<b>English-origin Nouns in HERITAGE JAPANESE</b>					
<b>Subject</b>	<b>Longer</b>			<b>Shorter</b>		
Correct mean	.281			.174		
Total N	36/128			23/132		
	FW	%	Token N	FW	%	Token N
<b>FOCUS PARTICLE</b>						
Presence	.84	66.7	4/6	.95	77.8	7/9
Absence	.48	26.2	32/122	.45	13.0	16/123
<b>Range</b>	<b>34</b>			<b>50</b>		
<b>VERBAL ADJACENCY</b>						
Adjacent	[.55]	32.9	25/76	[.51]	18.4	18/98
Non-adjacent	[.41]	20.9	9/43	[.47]	16.1	5/31
<b>Range</b>						
<b>SENTENCE-FINAL PARTICLE</b>						
Presence	[.50]	30.2	19/63	[.55]	19.6	9/46
Absence	[.50]	26.6	17/64	[.45]	16.3	14/86
<b>Range</b>						
<b>CLAUSE TYPE</b>						
Main	[.50]	28.2	22/78	[.53]	19.3	11/57
Subordinate	[.51]	28.6	14/49	[.48]	16.0	12/75
<b>Range</b>						

Next, Table 5.11 presents the results of the multivariate analysis of null marking on Japanese subject nouns for homeland Japanese speakers according to age. Again, only the focus particle factor group contributes a strong effect on the variant choice with a range value of 35 for younger speakers and 52 for older speakers. Null marking is favoured when a focus particle is present. None of the remaining factor groups has a significant effect on variant choice; however, a similar direction of effect is observed for each factor across the cohorts, with the exception of the neutralized effects for clause type among younger speakers. The detailed comparative

analysis according to age in homeland Japanese suggests that younger speakers and older speakers share the same underlying grammar for variable case marking on subject nouns, with a strong effect on null marking exerted by the presence of a focus particle. No ongoing change with regard to the linguistic conditioning of variation is found in homeland Japanese when the constraint hierarchies for younger speakers are compared with those of their older counterparts.

Table 5.11: Multivariate analysis of the contribution of linguistic factors selected as significant to the probability of null marking on Japanese subject nouns according to age for homeland Japanese speakers; factor groups not selected as significant are in square brackets.

	<b>Japanese Nouns in HOMELAND JAPANESE</b>					
<b>Subject</b>	<b>Younger</b>			<b>Older</b>		
Correct mean	.323			.178		
Total N	64/198			35/197		
	FW	%	Token N	FW	%	Token N
<b>FOCUS PARTICLE</b>						
Presence	.81	66.7	12/18	.98	88.9	8/9
Absence	.46	28.9	52/180	.46	14.4	27/188
<i>Range</i>	<b>35</b>			<b>52</b>		
<b>VERBAL ADJACENCY</b>						
Adjacent	[.52]	34.8	54/155	[.56]	19.0	24/126
Non-adjacent	[.41]	25.0	10/40	[.38]	17.2	11/64
<i>Range</i>						
<b>SENTENCE-FINAL PARTICLE</b>						
Presence	[.55]	36.7	36/98	[.60]	21.4	24/112
Absence	[.45]	27.3	27/99	[.37]	12.9	11/85
<i>Range</i>						
<b>CLAUSE TYPE</b>						
Main	[.50]	34.1	45/132	[.54]	19.8	22/111
Subordinate	[.50]	28.1	18/64	[.45]	15.1	13/86
<i>Range</i>						

### 5.7.2.2 Variable Case Marking on Nouns in Direct Object Position

In this section, we will investigate the variable case marking on direct object nouns. Table 5.12 displays the overall results of the multivariate analysis of the influence of social factors on the null marking of direct object nouns: English nouns and Japanese nouns in heritage Japanese, and Japanese nouns and loanwords in homeland Japanese. As indicated by the corrected mean, the overall rate of null marking on direct object nouns is higher in homeland Japanese (.536 for Japanese nouns and .618 for loanwords) than in the heritage variety (.462 for Japanese nouns and .495 for English-origin nouns). In homeland Japanese, age is again selected as a significant factor on variant choice for Japanese nouns with a range value of 12; younger speakers favour null marking with a probability of .56. Although the age factor is not selected as significant for loanwords, a similar direction of effect is observed. The factor of sex is not selected as significant. In the heritage variety, on the other hand, none of the social factors are selected as significant on variant choice for either the Japanese nouns or the English-origin nouns.

Table 5.12: Multivariate analysis of the contribution of social factors selected as significant to the probability of null marking on direct object nouns; factor groups not selected as significant are in square brackets.

	HERITAGE JAPANESE						HOMELAND JAPANESE					
Direct Object	English-origin nouns			Japanese nouns			Japanese nouns			Loanwords		
Corrected mean	.495			.465			.536			.618		
Total N	149/301			177/381			207/386			84/136		
	FW	%	Token N	FW	%	Token N	FW	%	Token N	FW	%	Token N
<b>LENGTH OF STAY IN CANADA</b>												
Longer	[.54]	53.1	76/143	[.52]	48.5	94/194	N/A			N/A		
Shorter	[.47]	46.2	73/158	[.48]	43.9	82/187	N/A			N/A		
Range												
<b>AGE</b>												
Younger	N/A			N/A			.56	59.7	111/186	[.55]	65.8	50/76
Older	N/A			N/A			.44	48.0	96/200	[.44]	56.7	34/60
Range							12					
<b>SEX</b>												
Female	[.55]	54.5	85/156	[.55]	51.1	95/186	[.55]	56.7	106/187	[.52]	62.5	45/72
Male	[.45]	44.1	64/145	[.45]	41.5	81/195	[.45]	50.8	101/199	[.47]	60.9	39/64
Range												

As for linguistic conditioning, the same factor groups as those operationalized for subject nouns are examined for direct object nouns. Table 5.13 synthesizes the results of the multivariate analysis of linguistic factors for nouns in direct object position. Inspection of the results for Japanese nouns in the homeland variety shows that the presence of a focus particle contributes the strongest effect to variant choice, as assessed by the range value of 41, followed by the factor of verbal adjacency with a range of 19. For Japanese nouns, null marking is favoured by the presence of a focus particle with a probability of .88 and when the noun is directly adjacent to the predicate with a probability of .52. These findings are generally aligned with those reported in the literature (Fry 2003; Fujii and Ono 2000; Matsuda 1995; Saito 1983; Tsutsui 1984). For loanwords, a significant effect on variant choice is returned for presence of a focus particle with

a range value of 46, followed by verbal adjacency with a range value of 44. The sentence-final particle factor group and the clause type factor group are both selected as non-significant. Although the same direction of effect is seen for clause type between Japanese nouns and loanwords, a neutralized effect is found for loanwords with regard to the sentence-final particle factor group.

Turning to the results for Japanese nouns in heritage Japanese, the focus particle factor group also contributes the strongest effect to variant choice, as indicated by the range value of 40, with the presence of a focus particle favouring null marking with a probability of .87. This factor exhibits a strong effect on variant choice for English-origin nouns with a range value of 51; null marking is favoured by the presence of a focus particle with a probability of .95. For both types of noun, the factor of verbal adjacency is selected as significant with a range value of 17 for Japanese nouns and with a range value of 31 for English-origin nouns. Null marking is favoured when the noun is adjacent to the predicate without any intervening material with a probability of .54 for Japanese nouns and .55 for English-origin nouns. For English-origin nouns, the sentence-final particle factor group is selected as significant with a range value of 15; null marking is favoured by the presence of a sentence-final particle with a probability of 61. This factor effect is neutralized for Japanese nouns. Clause type seems to have no appreciable effect on variant choice in heritage Japanese.

Table 5.13: Multivariate analysis of the contribution of linguistic factors selected as significant to the probability of null marking on direct object nouns; factor groups not selected as significant are in square brackets.

	HERITAGE JAPANESE						HOMELAND JAPANESE					
Direct Object	English-origin nouns			Japanese nouns			Japanese nouns			Loanwords		
Corrected mean	.495			.462			.536			.618		
Total N	149/301			176/381			207/386			68/115		
	FW	%	Token N	FW	%	Token N	FW	%	Token N	FW	%	Token N
<b>FOCUS PARTICLE</b>												
Presence	.95	95.5	21/22	.87	83.8	20/24	.88	89.3	25/28	.92	85.7	6/7
Absence	.44	45.9	128/279	.47	43.7	156/357	.47	50.8	182/358	.46	57.4	62/108
Range	<b>51</b>			<b>40</b>			<b>41</b>			<b>46</b>		
<b>VERBAL ADJACENCY</b>												
Adjacent	.55	54.6	136/249	.54	48.8	145/297	.52	55.7	190/341	.57	63.6	63/99
Non-adjacent	.24	25.5	12/47	.37	35.8	29/81	.33	38.1	16/42	.13	26.7	4/15
Range	<b>31</b>			<b>17</b>			<b>19</b>			<b>44</b>		
<b>SENTENCE-FINAL PARTICLE</b>												
Presence	.61	57.5	50/87	[.51]	47.4	64/135	[.57]	58.3	88/151	[.50]	58.8	30/51
Absence	.46	46.0	98/213	[.50]	45.0	109/242	[.46]	50.6	119/235	[.50]	58.7	37/63
Range	<b>15</b>											
<b>CLAUSE TYPE</b>												
Main	[.51]	50.0	51/102	[.52]	47.6	90/189	[.47]	52.5	96/183	[.47]	54.5	24/44
Subordinate	[.50]	49.2	98/199	[.48]	45.5	86/189	[.54]	53.8	105/195	[.52]	61.4	43/70
Range												

Comparing variable case marking patterns on direct object nouns between heritage Japanese and the homeland benchmark variety shows that the same underlying grammars are shared, particularly with regard to the factors of focus particle and verbal adjacency. A similar direction of effect is found for sentence-final particle even though the factor is selected as significant for English-origin nouns only. The only slight difference is seen for the non-significant factor of clause type in all types of direct object nouns where the direction of effect is reversed when the homeland variety is compared with the heritage variety in relation to this

factor group. However, consideration of the various lines of evidence emerging from the regression analysis suggests that the two varieties of Japanese share a similar underlying grammar for variable case marking on direct object nouns. Thus, no compelling evidence in favour of contact-induced change emerges from these analyses.

Lastly, we inquire into a possible ongoing language change in homeland Japanese with regard to variable case marking on Japanese direct object nouns since age is selected as a significant factor on variant choice. Table 5.14 displays the results of the multivariate analysis of linguistic factors for null marking on Japanese nouns in direct object position in homeland Japanese by age. Among older speakers, the focus particle factor group returns a knockout value, (i.e. null marking is categorically selected when a focus particle is present). On the other hand, this factor is not selected as significant among younger speakers. However, the same direction of effect is observed; null marking is more likely to occur with the presence of a focus particle (i.e. 80% of the time). The verbal adjacency factor group contributes an effect to variant choice among younger speakers with a range value of 34. Null marking is weakly favoured when the noun is adjacent to the predicate with a probability of .53. Although verbal adjacency does not have a significant effect on variant choice among older speakers, the same direction of effect is exhibited. Two other factors, sentence-final particle and clause type, are not selected as significant; the same direction of effect is found for the former while a reverse effect is observed for the latter factor.

Summarizing, when the environmental constraints are compared across age cohorts, there is a difference in relative strength of factor effects (i.e. focus particle and verbal adjacency). It may hint at possible language change in progress with regard to variable case marking of

Japanese direct object nouns in homeland Japanese. However, additional real-time data would be needed to confirm any evidence of ongoing change.

Table 5.14: Multivariate analysis of the contribution of linguistic factors selected as significant to the probability of null marking on Japanese direct object nouns according to age for homeland Japanese speakers; factor groups not selected as significant are in square brackets.

	<b>Japanese Nouns in HOMELAND JAPANESE</b>					
<b>Direct Object</b>	<b>Younger</b>			<b>Older</b>		
Corrected mean	.597			.480		
Total N	111/186			96/200		
	FW	%	Token N	FW	%	Token N
<b>FOCUS PARTICLE</b>						
Presence	[.74]	80.0	12/15	K.O	100.0	13/13
Absence	[.48]	57.9	99/171		44.4	83/187
<i>Range</i>						
<b>VERBAL ADJACENCY</b>						
Adjacent	.53	61.8	107/173	[.51]	49.4	83/168
Non-adjacent	.19	25.0	3/12	[.45]	43.3	13/30
<i>Range</i>	<b>34</b>					
<b>SENTENCE-FINAL PARTICLE</b>						
Presence	[.59]	63.0	51/81	[.56]	52.9	37/70
Absence	[.43]	57.1	60/105	[.47]	45.4	59/130
<i>Range</i>						
<b>CLAUSE TYPE</b>						
Main	[.42]	55.5	61/110	[.51]	47.9	35/73
Subordinate	[.63]	66.7	48/72	[.49]	46.3	57/123
<i>Range</i>						

## 5.8 Discussion

In this chapter, we investigated the underlying conditioning of variable case marking of subject and direct object nouns in order to assess contact-induced structural change in heritage Japanese. Variable case marking patterns in heritage Japanese were compared systematically and quantitatively with their homeland counterpart.

Variable case marking is one of the linguistic variables that has been studied to better understand contact effects. The present study compared the rates of null marking on Japanese nouns and English origin-nouns in heritage Japanese with those on Japanese nouns and loanwords in the homeland benchmark variety. It was found that heritage Japanese speakers show less null marking than homeland Japanese speakers. At first glance, a difference in the quantitative patterning of null marking in the two varieties may indicate that heritage Japanese speakers are diverging from the homeland Japanese due to contact with English. However, as already discussed, rates can fluctuate for all kinds of non-linguistic reasons (Tagliamonte and Poplack 2001; Torres-Cacoulllos and Travis 2010). Here we have attempted to transcend the traditional reliance on rates by emphasizing the importance of comparing the underlying structure of variation. A summary of the effects uncovered by this comparative endeavor is presented in Table 5.15.

Table 5.15: Summary of findings on the variable case marking between varieties of Japanese

	SUBJECT				DIRECT OBJECT			
	HERITAGE JAPANESE		HOMELAND JAPANESE		HERITAGE JAPANESE		HOMELAND JAPANESE	
Favouring Null marking	English-origin nouns	Japanese nouns	Japanese nouns	Loanwords	English-origin nouns	Japanese nouns	Japanese nouns	Loanwords
Focus particle (Presence)	✓	✓	✓	✓	✓	✓	✓	✓
Adjacency to the predicate	✓	✓	✓	×	✓	✓	✓	✓
Sentence-final particle (Presence)	✓	✓	✓	✓	✓	—	✓	—
Main Clause	✓	✓	✓	✓	—	✓	×	×
✓	Significant with the same constraint hierarchy							
✓	Not significant but showing the same direction of effect							
—	Neutralized effect							
×	Not significant but showing the opposite direction of effect							

Looking at case marking on subject nouns, all types of nouns in heritage Japanese and homeland Japanese share a similar direction of effect for all independent variables, except for verbal adjacency for loanwords in homeland Japanese. For variable case marking on direct object nouns, by and large, the same direction of effect is shared for all types of nouns between heritage Japanese and homeland Japanese. The only anomaly is observed for clause type; however, it is not selected as a significant factor in relation to variant choice.

Considering the multiple lines of evidence revealed by the multivariate analysis, we unearthed little compelling evidence that the underlying grammar for variable case marking in heritage Japanese fundamentally diverges from the underlying grammar found in the homeland benchmark variety.

## CHAPTER 6

### CONCLUSION

The present study set out to investigate whether or not heritage Japanese, as spoken in Canada, is showing structural change due to contact with English. An important motivation for this study came from claims in the literature that prolonged and sustained contact between a minority (heritage) language and a majority language (in this case, English) provides the ideal social conditions for linguistic convergence. According to much of the literature on first language (L1) loss or attrition (see e.g. Seliger and Vago 1991), diminished opportunities for heritage language use arising from lack of regular (or restricted) contact with other speakers of the heritage language may lead to attrition or erosion of the structure of the first language (Seliger and Vago 1991:3). Lack of regular exposure to speakers of a heritage variety, coupled with the acquisition of increasing fluency in the language of the host community (the L2), are widely believed to be the cause of L2 structural influence on the grammar of the heritage language. External influence from the L2 manifests itself most obviously in the form of lexical borrowings, but it is also acknowledged in the literature that this influence may be more subtle and may manifest itself in the gradual loss, modification of, or increase in the use of an existing grammatical feature in the heritage language under the influence of the L2 language. In addition, it has been claimed that grammatical properties found at the syntax/semantics/pragmatics interface are vulnerable to convergence (Bullock and Toribio 2004, Sorace 2011).

By employing an empirically accountable variationist sociolinguistic approach, I first located *conflict sites* between Japanese and English (Poplack and Meechan 1998) and selected

specific linguistic variables as diagnostics of possible contact-induced change. The three variables targeted in the present study, namely, *variable realization of subject pronouns*, *variable case marking of subject nouns*, and *variable case marking of direct object nouns*, are ideal linguistic features to examine in the speech of Japanese-English bilinguals because both linguistic variables straddle syntactic as well as discourse-pragmatic components of the grammar. In the literature on language contact, languages that license variable pronominal realization have been viewed as susceptible to change when in sustained contact with English, a non-pro-drop language (Lapidus and Otheguy 2005; Otheguy et al. 2007). Similar susceptibility to change is also reported in the instance of case marking in heritage languages such as Russian, Polish, and German (see e.g. Clyne 2003; Huffines 1989, 1999; Polinsky 1995, 1997; Preston 1986). When the heritage speakers of these languages are compared with speakers of their respective homeland varieties, their usage has been found to deviate significantly from monolingual baseline norms. In the case of contact with English, the inference that case marking in a heritage variety is subject to erosion, or change, is motivated by the fact that in English, there is no parallel system requiring overt marking of subject and direct objects. These claims generated a number of hypotheses that I put to the test in the present study.

The key methodological component of the present study involved the combined use of variationist sociolinguistics with multiple comparison points to pursue the inference of contact-induced change. Spontaneous speech production of Japanese-English bilinguals, obtained from the *Corpus of Japanese-English Contact in Canada*, was systematically compared with the vernacular speech data of an appropriate homeland vernacular variety (the *Corpus of Kwansai Spoken Japanese*, Heffernan 2012). I attached particular importance to the comparison of heritage Japanese with a *vernacular* non-contact control variety on the grounds that comparison

with a standard or idealized version of Japanese would be inappropriate (See Dorian 1993; Mougeon et al. 2005; Poplack and Levey 2010).

A further reason for the importance attached in this study to the analysis of natural speech data derives from the observation that much previous research on heritage speech varieties has relied on assorted methodologies, sometimes involving cherry-picked examples, decontextualized data, introspective evaluation of language data, or grammaticality judgements. While these alternative methodologies may be informative, our emphasis on natural speech data is justified by the fact that if change is taking place, appropriate evidence is optimally located in vernacular speech, construed by sociolinguists to yield the most systematic data for analysis (Labov 1984). Furthermore, an intrinsic property of speech, and one which lies at the heart of the conceptual and methodological framework characterizing this study, is the notion of *inherent variability*. The underlying structure of this variability constitutes the most important basis for our comparative analysis. Observing that many claims of contact-induced change are exclusively based on rates of linguistic features alone, this study attached particular importance to constraint hierarchies generated by a multivariate analysis of spontaneous speech data. It is these constraint hierarchies that formed the cornerstone of the quantitative analysis presented in this study and offered “the most penetrating characterization of variable structure” (Erker and Guy 2012:546). Where the constraint hierarchies exhibited non-trivial parallels across the heritage variety and the homeland baseline, it was inferred that there are substantial structural affinities between the comparison varieties for the variable in question.

Drawing on this methodological framework, we are now in a position to review the research objectives posed at the outset of this study, and evaluate them in the light of the information that was uncovered in the present study. The first research objective in this study

was to investigate the major sociolinguistic patterns of variation in the speech of bilingual speakers in heritage Japanese. In terms of the first variable analyzed, *variable realization of subject pronouns*, it was found that in heritage Japanese, this variable is conditioned by the factor group of *subject continuity* (i.e. switch reference); the same subject referent favoured and a different subject referent disfavoured the null variant. It was also observed that the null variant was favoured by the second person pronoun and disfavoured by the first person pronoun according to the effect of *grammatical person*. None of the social factors, such as speakers' length of stay in Canada and sex, was significant for variant choice. Although in the literature (Thomason 2001; Thomason and Kaufman 1988; Winford 2003), the duration of contact with the majority language is generally considered to have an impact on contact-induced change, our results indicate that there is no such contact effect on heritage Japanese spoken in Canada when examining the effect of speakers' length of stay on usage patterns in Japanese.

The next variables I examined were *variable case marking on subject nouns* and *variable case marking on direct object nouns*. According to the literature on this topic, case marking is a variable rather than an obligatory process in colloquial speech (e.g. Martin 1975; Matsuda 1995; Shibamoto 1985; Shibatani 1990; Takano 1998); similarly, our findings confirm that in heritage Japanese and homeland Japanese case marking is also variable, comprising either null marking or overt marking. As for variable subject case marking on English-origin nouns and Japanese nouns in heritage Japanese, I found that null marking is favoured: (a) when a subject noun occurs with a focus particle; and (b) when a subject Japanese noun is adjacent to a predicate, with the same direction of effect instantiated for English-origin nouns. The factor groups *sentence-final particle* and *clause type* were found to be non-significant for heritage Japanese, but their respective direction of effect is parallel. As for variable direct object case marking in heritage

Japanese, it was found that null marking is favoured: (a) by the presence of a focus particle; and (b) when the direct object is adjacent to the predicate. In addition, null marking is favoured by the presence of a sentence-final particle for English-origin nouns, while the same effect is neutralized for Japanese nouns in the heritage speaker data set. Clause type was not selected as significant. With regard to the social conditioning of variant choice, speakers' length of stay in Canada shows a significant effect for English-origin subject nouns only; speakers who have resided in Canada for a longer period favour null marking. Taking into account previous claims relating to the effects of length of contact on language change (Thomason 2001; Thomason and Kaufman 1988; Winford 2003), it was initially assumed that there might be some differences in the underlying grammars based on the speakers' length of stay in Canada. However, it was found that heritage Japanese speakers who have resided in Canada for a longer period share the same underlying grammar for subject marking on English-origin nouns with those speakers who have resided in Canada for a shorter period. Although the length of stay in Canada was not selected as a significant factor for variant choice for Japanese nouns in subject position, and English-origin and Japanese nouns in direct object position, the same direction of effect was found in terms of length of stay: speakers who have been in Canada for a longer period are more likely to show null marking. No significant effect of sex was observed for either type of noun in heritage Japanese.

The second and most important objective was to assess whether the resultant social and linguistic patterns of variability instantiated by heritage speakers of Japanese could be explained by invoking English influence. Systematic comparison of the underlying patterns of variability in heritage Japanese with the homeland benchmark vernacular variety turned up a number of crucial findings.

First, it was found that the underlying grammar for the variable realization of subject pronouns in heritage Japanese is shared by the homeland benchmark variety, showing the same significant effects and constraint hierarchies for grammatical person and subject continuity, although there is a slight difference in the ranking of statistically significant factor groups according to their relative strength of effect in the two comparison groups.

As for variable case marking of subject and direct object nouns, the constraint hierarchies in heritage Japanese were identical with those in the homeland variety for focus particle, selected as statistically significant. Presence of a focus particle favours null marking consistently for all types of nouns in both heritage Japanese and homeland Japanese. The constraint hierarchies (and direction of the effect) for the other significant factor groups of verbal adjacency and sentence-final particle were identical between heritage Japanese and the homeland variety, with the exception of (a) a reversed direction of effect for loanword subject nouns in heritage Japanese for the non-significant factor group of verbal adjacency, and (b) a neutralized effect for Japanese nouns in heritage Japanese and loanwords in homeland Japanese when these nouns are located in direct object position. In both varieties of Japanese, clause type was not selected as a significant factor influencing variant choice for any type of noun. The direction of effect for this factor group is the same for all subject nouns in heritage and homeland Japanese, but the direction of effect is reversed for direct object nouns when the two varieties are compared. Considered in the aggregate, it was found that the constraint hierarchies exhibit a number of parallels across comparison varieties. This finding bolsters the general conclusion that there is little evidence indicating that extensive contact with English has had any discernible impact on structural patterns in this sector of the heritage grammar. Simplification or loss of the case marking system has been widely reported in the literature for other heritage languages in contact with English

(e.g. Clyne 2003; Larmouth 1974; Meyerstein 1969; Preston 1986). However, no such simplification or loss was found in heritage Japanese in contact with English in Canada.

Prior to the dissertation, lone English-origin subject nouns and direct object nouns in the same heritage Japanese data were identified as borrowing and distinguished clearly from codeswitched items, by comparing their underlying case marking system with that of their native Japanese counterparts (Yoshizumi 2012). It was found that lone English-origin items in otherwise Japanese discourse were integrated into Japanese grammar. In the present study, the case marking system in heritage Japanese was further compared with that of homeland Japanese on the subject of nouns and loanwords in the monolingual homeland variety. From the finding that heritage Japanese and homeland Japanese share a similar underlying grammar with regard to case marking, the status of lone English-origin nouns as borrowed items is further corroborated.

In addition to the above findings, our results indicate that younger speakers of homeland Japanese favour null marking on Japanese nouns in both the subject and the direct object position. Comparison of the underlying grammars for younger and older speakers shows that constraint hierarchies (and their respective direction of effect) for significant factors (i.e., presence of a focus particle and verbal adjacency) were identical for the two cohorts. For non-significant factors, the same direction of effect was found for sentence-final particle and clause type for subject nouns. The same direction of effect was also found for sentence-final particles in direct object position, while a reverse direction is observed for clause type (although the respective effects seem to be neutralized for older speakers). In sum, variable case marking on Japanese nouns is linguistically conditioned similarly when younger speakers and older speakers in the homeland data set are compared. The apparent time differences in rates of case marker ellipsis and in relative strength of factor effects found in the analysis merit additional investigation

drawing on real-time data in order to verify whether the patterns uncovered are indicative of possible linguistic change in progress.

Summarizing, the multiple lines of evidence emerging from the empirical quantitative analyses of the variables targeted in this thesis converge in indicating that heritage Japanese, as spoken in Canada, broadly shares the same underlying grammar as homeland Japanese. Structural affinities in variable patterning shared by heritage and homeland varieties reveal little compelling evidence indicating that heritage Japanese exhibits structural change due to contact with English.

Finally, our observations emerging from the present study warrant some discussion relating to the issues foregrounded at the start of this dissertation. First, it has been claimed that typological incongruence tends to diminish opportunities for structural convergence (Thomason 2001). When it comes to heritage Japanese in contact with English, that seems to be the case. In heritage Japanese, no clear divergence from homeland Japanese was found; moreover, no convergence with English was confirmed, particularly in relation to variable realization of subject pronouns.

However, recall the claim that language change or language attrition is more likely to occur in the domain that is situated at the interface between (morpho-)syntax and discourse (Bullock and Toribio 2004; Sorace 2005, 2011). The linguistic variables examined in the present study were indeed in the domains involving the interface. It was assumed that contact-induced change could happen; however, as our findings demonstrated, no substantial evidence of contact-induced change was found in these sectors of Japanese grammar: variable realization of subject pronouns and variable case marking on subject and direct object nouns.

Furthermore, studies on attrition raise the issue that linguistic variability observed in heritage languages is often associated with language change or attrition without discussing whether or not the observed variation constitutes variability inherent in the language. From the comparative analysis of heritage Japanese and homeland Japanese in the present study, the observed variation for realization of subject pronouns and case marking in heritage Japanese do not support the possible language attrition among heritage Japanese speakers in Canada. Rather, it is demonstrated that their underlying grammar in those linguistic sectors is shared by the monolingual speakers in homeland Japanese, indicating that heritage Japanese speakers in Canada are not undergoing language attrition or language change but preserving the inherent variability as found in homeland Japanese.

Finally, following the hypothesis that structural language change may be associated with a long period of contact between languages (Thomason 2001; Thomason and Kaufman 1988; Winford 2003), heritage speakers' length of residence in Canada, as a proxy of their length of contact period with English, was investigated. However, the present study did not find any effect of speaker's length of residence in Canada on contact-induced structural change. In fact, several studies have reported no influence of the temporal interval since emigration on language attrition (de Bot and Clyne 1994; de Bot et al. 1991). De Bot and Clyne (1994:17), who investigated language attrition of German speakers in Australia, further states that "first-language attrition does not necessarily take place in an immigrant setting." This would be the case for heritage Japanese speakers in Canada.

I acknowledge that although the heritage speakers in this study are in a Japanese-English bilingual setting (i.e. at work and at home), the speakers' language attitudes, level of bilingual ability, language choice and so on, were not considered in detail in the present analysis. An

investigation of the correlation between these factors and variable linguistic patterns may offer additional insights into Japanese-English contact in Canada. In addition, the bilingual speakers of Japanese and English for this study were limited to those who were born in Japan and immigrated to Canada in their early adulthood. Further research based on data from Japanese-English child bilinguals growing up in a bilingual environment, and whose linguistic systems, by virtue of their age, are still malleable, may provide additional insights into Japanese-English contact. For example, is it the case that Japanese-English bilingual children, who are still in the process of acquiring both languages, are more susceptible to contact-induced change than their adult caregivers? This is a fruitful line of inquiry to be pursued in future research.

In conclusion, the present study has endeavoured to make an important contribution to understanding aspects of Japanese-English bilingualism in Canada. Although this study did not turn up substantial evidence of contact-induced change in heritage Japanese with respect to the variables that were targeted for analysis, it cannot be ruled out that contact-induced change has not taken place in other areas of speakers' grammar. Notwithstanding this possibility, it is important to stress that, in line with Thomason's (2001:93) observation that contact-induced change is unlikely to be restricted to isolated incidents of structural interference, this study deliberately targeted three variables that are claimed to be susceptible to change in contact scenarios. Only more accountable analyses of the type pursued in this study will confirm whether or not change is found in other grammatical domains in the speech of Japanese-English bilinguals. To conclude, the findings of the present study dovetail with those reported in other studies of language contact (e.g. Poplack et al. 2012a) in demonstrating that inspection of the fine-grained conditioning of variation is a methodological prerequisite for scientifically assessing hypotheses of contact-induced change.

## BIBLIOGRAPHY

- Adalar, Nevin & Tagliamonte, Sali. 1998. Borrowed nouns; bilingual people: The case of the "Londrali" in Northern Cyprus. *International Journal of Bilingualism* 2(2): 139-160.
- Altenberg, Evelyin P. 1991. Assessing first language vulnerability to attrition. In H.W. Seliger & R. M. Vago (eds.), *First Language Attrition*. Cambridge: Cambridge University Press. 189-206.
- Appel, René & Muysken, Pieter. 1987. *Language Contact and Bilingualism*. London: Edward Arnold.
- Azuma, Shoji. 1987. Conversational code-switching in Japanese/English: a study on Japanese Americans. *The Japan Association of College English Teachers Bulletin* 18: 1-19.
- Azuma, Shoji. 1993. The frame-content hypothesis in speech production: evidence from intrasentential code-switching. *Linguistics* 31: 1071-1093.
- Azuma, Shoji. 1997. Lexical categories and code-switching: A study of Japanese/English code-switching in Japan. *Journal of the Association of Teachers of Japanese* 31(2): 1-24.
- Backhaus, Peter. 2006. Multilingualism in Tokyo: A look into the linguistic landscape. *International Journal of Multilingualism* 3(1): 52-66.
- Backus, Ad. 2005. Code-switching and language change: One thing leads to another? *International Journal of Bilingualism* 9(3-4): 307-340.
- Bailey, Robert. 2002. The quantitative paradigm. In J. K. Chambers, P. Trudgill & N. Schilling-Estes (eds.), *The Handbook of Language Variation and Change*. Oxford: Blackwell. 117-141.
- Baldi, Philip. 1990. Introduction: The comparative method. In P. Baldi (ed.), *Linguistic Change and Reconstruction Methodology*. Berlin: Mouton de Gruyter. 1-13.
- Barke, Andrew J. 2000. What 'you' can say in Japanese: Gender as a pragmatic factor in Japanese second-person reference through history. In Janet Holmes (ed.), *Gendered speech in social context: perspectives from gown and town*. Wellington: Victoria University Press. 63-73.
- Bayley, Robert & Pease-Alvarez, Lucinda. 1997. Null pronoun variation in Mexican-descent children's narrative discourse. *Language Variation and Change* 9:349-371.

- Bentahila, Abdelâli & Davies, Eirlys E. 1991. Constraints on code-switching: A look beyond grammar. In *Papers for the symposium on code-switching in bilingual studies: Theory, significance and perspectives*. Strasbourg: European Science Foundation. 369-404.
- Bentahila, Abdelâli & Davies, Eirlys E. 1995. Patterns of code-switching and patterns of language contact. *Lingua* 96(2-3): 75-93.
- Bentahila, Abdelâli & Davies, Eirlys E. 1998. Codeswitching: An equal partnership? In R. Jacobson (ed.), *Codeswitching Worldwide*. Berlin: Mouton de Gruyter. 25-49.
- Biberauer, Theresa, Holmberg, Andres, Roberts, Ian & Sheehan, Michelle. 2010. *Parametric Variation: Null Subjects in Minimalist Theory*. New York: Cambridge University Press.
- Boyd, Sally. 1993. Attrition or expansion? Changes in the lexicon of Finnish and American adult bilinguals in Sweden. In K. Hyltenstam & Å. Viberg (eds.), *Progression and Regression in Language: Sociocultural, Neurophysical, and Linguistic Perspectives*. Cambridge: Cambridge University Press. 386-411.
- Boztepe, Erman. 2003. Issues in code-switching: Competing theories and models. *Working Papers in TESOL and Applied Linguistics* 3 (2): 1-27.
- Budzhak-Jones, Svitlana. 1998. Against word-internal codeswitching: Evidence from Ukrainian-English bilingualism. *International Journal of Bilingualism* 2(2): 161-182.
- Budzhak-Jones, Svitlana & Poplack, Shana. 1997. Two generations, two strategies: The fate of bare English-origin nouns in Ukrainian. *Journal of Sociolinguistics* 1(2): 225-258.
- Bullock, Barbara E. 2009. Phonetic reflexes of code-switching. In B. Bullock & J. Toribio (eds.), *Linguistic Code-switching*. Cambridge: Cambridge University Press. 163-181.
- Bullock, Barbara E. & Gerfen, Chip. 2004. Phonological convergence in a contracting language variety. *Bilingualism: Language and Cognition* 7(2): 95-104.
- Bullock, Barbara E. & Toribio, Jacqueline A. 2004. Convergence as an emergent property in bilingual speech. *Bilingualism: Language and Cognition* 7(2): 91-93.
- Bynon, Theodora. 1977. *Historical Linguistics*. Cambridge: Cambridge University Press.
- Camacho, José A. 2013. *Null Subjects*. Cambridge: Cambridge University Press.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris
- Chomsky, Noam & Lasnik, Howard. 1977. Filters and control. *Linguistic Inquiry* 8: 425-504.

- Clancy, Patricia. 1980. Referential choice in English and Japanese narrative discourse. In Wallace Chafe (ed.), *The Pear Stories: Cognitive and Linguistic Aspect of Narrative Production*. Norwood, NJ: Ablex. 127-202.
- Clyne, Michael. 2003. *Dynamics of Language Contact*. Cambridge: Cambridge University Press.
- Corrigan, Karen P. 2010. Language contact and grammatical theory. In R. Hickey (ed.), *The Handbook of Language Contact*. 106-127.
- Cummins, Jim. 2005. A proposal for action: strategies for recognizing heritage language competence as a learning resource within the mainstream classroom. *Modern Language Journal* 89: 585-592.
- Daulton, Frank E. 2008. *Japan's Built-in Lexicon of English-based Loanwords*. Clevedon /Philadelphia: Multilingual Matters.
- De Bot, Kess & Clyne, Michael. 1994. A 16-year longitudinal study of language attrition in Dutch immigrants in Australia. *Journal of Multilingual and Multicultural Development* 15 (1): 17-28.
- De Bot, Kess, Gommans, Paul & Rossing, Carola. 1991. L1 loss in an L2 environment: Dutch immigrants in France. In H. W. Sc liger & R. M.Vago (eds.), *First Language Attrition*. Cambridge: Cambridge University Press. 87-98.
- Di Sciullo, Anne-Marie, Muysken, Pieter & Singh, Rajendra. 1986. Government and code-mixing. *Journal of Linguistics* 22: 1-24
- Dorian, Nancy. 1981. *Language death: The life cycle of a Scottish Gaelic Dialect*. Philadelphia: University of Pennsylvania Press.
- Edwards, Malcolm & Gardner-Chloros, Penelope. 2007. Compound verbs in code-switching: bilinguals making do? *International Journal of Bilingualism* 11(1): 73-91.
- Erker, Daniel & Guy, Gregory R. 2012. The role of lexical frequency in syntactic variability: Variable subject person pronoun expression in Spanish. *Language* 88(3): 526-557.
- Eze, Ejike. 1997. *Aspects of language contact: A variationist perspective on codeswitching and borrowing in Igbo-English bilingual discourse*. Doctoral dissertation. University of Ottawa.
- Eze, Ejike. 1998. Lending credence to a borrowing analysis: Lone English-origin incorporations in Igbo discourse. *International Journal of Bilingualism* 2(2): 183-202.
- Flores-Ferrán, Nydia. 2004. Spanish subject personal pronoun use in New York City Puerto Ricans: Can we rest the case of English contact? *Language Variation and Change* 16: 49-73.

- Fotos, Sandra S. 1995. Japanese-English conversational codeswitching in balanced and limited proficiency bilinguals. *Journal of Multilingualism and Multiculturalism* 1 (1): 2-15.
- Fry, John. 2003. *Ellipsis and Wa-marking in Japanese Conversation*. New York: Routledge.
- Fujii, Noriko. 1991. *Historical Discourse Analysis: Grammatical Subjects in Japanese*. Berlin, New York: Mouton de Gruyter.
- Fujii, Noriko & Ono, Tsuyoshi. 2000. The occurrence and non-occurrence of the Japanese direct object marker *o* in conversation. *Studies in Language* 24(1): 1-39.
- Gardner-Chloros, Penelope. 2009. *Code-switching*. Cambridge: Cambridge University Press.
- Grimshaw, Jane & Mester, Armin. 1988. Light Verbs and  $\theta$ -Marking. *Linguistic Inquiry* 19: 205-232.
- Gumperz, John J. 1977. The sociolinguistic significance of conversational code-switching. *RELC Journal* 8(2): 1-34.
- Guy, Gregory R. 1988. Advanced VARBUL analysis. In K. Ferrara, B. Brown, K. Walters, and J. Baugh (eds.), *Linguistic Change and Contact*. Austin: Department of Linguistics, University of Texas at Austin. 124-136.
- Haegeman, Liliane. 1990. Understood subjects in English diaries. *Multilingua* 9(2): 157-99.
- Haegeman, Liliane. 1997. Register truncation and subject omission in English and in French. *English Language and Linguistics* 1: 233-270.
- Haegeman, Liliane. 2002. Non-overt subject pronouns in written English. In S. Scholz, M. Klages, E. Hantson & U. Römer (eds), *Language, Context and Cognition. Papers in Honour of Wolf Dietrich Bald's 60th Birthday*. München: Langenscheidt, Longman. 135-149.
- Haegeman, Liliane. 2013. The syntax of registers: diary subject omission and the privilege of the root. *Lingua* 130: 88-110.
- Haegeman, Liliane, & Ihsane, Tabea. 1999. Subject ellipsis in embedded clauses in English. *English Language and Linguistics* 3(1): 117-45.
- Haegeman, Liliane, & Ihsane, Tabea. 2001. Adult null subjects in the non-pro-drop languages: two diary dialects. *Language Acquisition* 9(4): 329-46.
- Harrison, Brian. 2000. *Passing on the Language: Heritage Language Diversity in Canada*. <http://www.statcan.gc.ca/pub/11-008-x/2000002/article/5165-eng.pdf>
- Harrmann, Harald. 1989. *Symbolic Values of Foreign Language Use: From the Japanese Case to a General Sociolinguistic Perspective*. Berlin: Mouton de Gruyter.

- Harvie, Dawn. 1998. Null subject in English: Wonder if it exists? *Cahiers Linguistique d'Ottawa* 26: 15–25.
- Hasegawa, Yoko. 2001. *The Routledge Course in Japanese Translation*. London: Routledge.
- Hasegawa, Yoko. 2014. *Japanese: A Linguistic Introduction*. Cambridge: Cambridge University Press.
- Haugen, Einar. 1950. The analysis of linguistic borrowing. *Language* 26: 210-231.
- Heffernan, Kevin. 2012. An introduction to the Corpus of Kansai Spoken Japanese. *Journal of Policy Studies* 41: 157-164.
- Heine, Bernd & Kuteva, Tania. 2005. *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.
- Heine, Bernd & Kuteva, Tania. 2010. Contact and grammaticalization. In R. Hickey (ed.), *The Handbook of Language Contact*. Oxford: Blackwell. 86-105.
- Hinds, John. 1978. Anaphora in Japanese conversation. In J. Hinds (ed.), *Anaphora in Discourse*. Alberta: Linguistic Research, Inc. 136-179.
- Hinds, John. 1980. Japanese conversation, discourse structure, and ellipsis. *Discourse Processes* 3(2): 263-286.
- Hinds, John. 1982. *Ellipsis in Japanese*. Edmonton: Linguistic Research, Inc.
- Hinds, John. 1983. Topic continuity in Japanese. In T. Givón (ed.), *Topic Continuity in Discourse: A Quantitative Cross-language Study*. Amsterdam: John Benjamins. 43-93.
- Hoenigswald, Henry. 1960. *Language Change and Linguistic Reconstruction*. Chicago: University of Chicago Press.
- Hoffer, Bates L., Beard, Grace Y. & Honna, Nobuyuki. 1983. Japanese use of English loans. *Papers in Linguistics* 16 (1-2): 35-56.
- Hoffer, Bates L. & Honna, Nobuyuki. 1988. The influx of English into the Japanese language. *Southwest Journal of Linguistics* 8(2): 15-33.
- Hooper, Joan. B. & Thompson, Sandra. A. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4: 465-497.
- Huang, James C.-T. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15(4): 531-574.

- Huffines, Marion L. 1989. Case usage among the Pennsylvania German sectarians and nonsectarians. In N. C. Dorian (ed.), *Investigating Obsolescence Studies in Language Contraction and Death*. Cambridge: Cambridge University Press. 211-226.
- Huffines, Marion L. 1991. Pennsylvania German: convergence and change as strategies of discourse. In H. W. Seliger & R. M. Vago (eds.), *First Language Attrition*. Cambridge: Cambridge University Press. 125-137.
- Iwasaki, Shoichi. 1993. *Subjectivity in Grammar and Discourse: Theoretical Considerations and a Case Study of Japanese Spoken Discourse*. Amsterdam: John Benjamins.
- Iwasaki, Shoichi. 2013. *Japanese*. Amsterdam: John Benjamins.
- Jaeggli, Osvaldo & Safir, Kenneth J. 1989. *The Null Subject Parameter*. Dordrecht: Foris.
- Johanson, Lars. 1993. Code-copying in immigrant Turkish. In G. Extra & L. Verhoeven (eds.), *Immigrant Languages in Europe*. Clevedon: Multilingual Matters. 197-221.
- Johanson, Lars. 2002. *Structural Factors in Turkic Language Contact*. Richmond: Curzon Press.
- Johnson, Yuki. 2008. *Fundamentals of Japanese Grammar: Comprehensive Acquisition*. Honolulu: University of Hawaii Press.
- Kageyama, Taroo. 1977. Incorporation and Sino-Japanese Verbs. *Papers in Japanese Linguistics* 5: 117-156.
- Kageyama, Taroo. 1991. Light verb constructions and the syntax-morphology Interface. In H. Nakajima (ed.), *Current English Linguistics in Japan*. Berlin: Mouton de Gruyter. 16-203.
- Kameyama, Megumi. 1986. *Zero anaphora: the case of Japanese*. Doctoral dissertation. Stanford University.
- Kayama, Yuhko. 2003. L1 Acquisition of Japanese zero pronouns: The effect of discourse factors. *Proceeding of the 2003 Annual Conference of the Canadian Linguistic Association*. 109-120.
- King, Ruth. 2000. *The Lexical Basis of Grammatical Borrowing: Prince Edward Island French Case Study*. Amsterdam/Philadelphia: John Benjamins.
- Kishimoto, Hideki. 2000. Locative verbs, agreement, and object shift in Japanese. *The Linguistic Review* 17:53-109.
- Koizumi, Masatoshi. 2008. Nominative object. In S. Miyagawa & M. Saito (eds.), *The Oxford Handbook of Japanese Linguistics*. Oxford: Oxford University Press. 141-164.
- Kuno, Susumu. 1973a. *The Structure of the Japanese Language*. Cambridge, Massachusetts: MIT Press.

- Kuno, Susumu. 1973b. *Nihon bunpokenkyu (Study of Japanese Grammar)*. Tokyo: Taishukan.
- Kuno, Susumu. 1978. *Danwa no bunpo (Grammar of Conversation)*. Tokyo: Taishukan.
- Kuroda, Shigeyuki. 1965. *Generative grammatical studies in the Japanese language*. Doctoral dissertation, Department of Linguistics and Philosophy, MIT
- Labov, William. 1963. The social motivation of a sound change. *Word* 19:273-309.
- Labov, William. 1966. *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov, William. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45:715-762.
- Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, William. 1984. Field methods of the project on linguistic change and variation. In J. Baugh & J. Sherzer (eds.), *Language in Use: Readings in Sociolinguistics*. Englewood Cliffs, NJ: Prentice Hall. 28-53.
- Labov, William. 1994. *Principles of Linguistic Change: Internal factors*. Oxford: Blackwell.
- Labov, William. 2001. *Principles of Linguistic Change: Social Factors*. Oxford: Blackwell.
- Laleko, Oksana & Polinsky, Maria. 2013. Marking topic or marking case: A comparative investigation of heritage Japanese and heritage Korean. *Heritage Language Journal* 10(2): 178-202.
- Laleko, Oksana & Kawamura, H. 2011. Grammatical restructuring in intergenerational language shift: A case of heritage Japanese. *Paper from the 56th Annual Conference of the International Linguistics Association*. Rutgers University, New Brunswick, NJ.
- Lapidus, Naomi & Otheguy, Ricardo. 2005. Overt nonspecific *Ellos* in Spanish in New York. *Spanish Context* 2: 157-174.
- Larmouth, Donald W. 1974. Differential interference in American Finnish cases. *Language* 50: 356-366.
- Lee, Duck-Young & Yonezawa, Yoko. 2008. The role of the overt expression of first and second person subject in Japanese. *Journal of Pragmatics* 40: 733-767.
- Leisiö, Larisa. 2006. Genitive subjects and objects in the speech of Finland Russians. *Journal of Slavic Linguistics* 14: 289-316.
- Loveday, Leo J. 1996. *Language Contact in Japan: A Sociolinguistic History*. Oxford: Oxford University Press.

- MacSwan, Jeff. 1999. *A Minimalist Approach to Intrasentential Code-Switching*. New York: Garland.
- MacSwan, Jeff. 2005. Codeswitching and generative grammar: A critique of the MLF model and some remarks on «modified minimalism». *Bilingualism: Language and Cognition* 8: 1-22.
- Martin, Samuel. E. 1975. *A Reference Grammar of Japanese*. London: Yale University Press.
- Masunaga, Kiyoko. 1988. Case deletion and discourse context. In W. Poser (ed.), *Papers from the Second International Workshop on Japanese Syntax*. CSLI Publications. 145-156.
- Matras, Yaron. 1998. Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36 (2): 281-331.
- Matras, Yaron. 2009. *Language Contact*. Cambridge: Cambridge University Press.
- Matsuda, Kenjiro. 1995. *Variable zero-marking of (o) in Tokyo Japanese*. Doctoral dissertation, University of Pennsylvania.
- Maynard, Senko. K. 1997. *Japanese Communication: Language and Thought in Context*. Honolulu: University of Hawai'i Press.
- Maynard, Senko. K. 2002. *Kaiwa bunseki (Conversation Analysis)*. Tokyo: Kuroshio.
- Meechan, Marjory & Poplack, Shana. 1995. Orphan categories in bilingual discourse: Adjectivization strategies in Wolof-French and Fongbe-French. *Language Variation and Change* 7: 169–194.
- Meyerstein, Goldie. P. 1969. Interference in prepositional phrases: Immigrant Slovak in America. *Lingua* 22: 63-80.
- Miller, Roy A. 1967. *The Japanese Language*. Chicago: University of Chicago Press.
- Milroy, James. 1992. *Linguistic Variation and Change*. Oxford: Blackwell.
- Milroy, Lesley & Gordon, Matthew. 2003. *Sociolinguistics: Method and Interpretation*. Oxford: Blackwell.
- Mitchell, Bruce & Robinson, Fred C. 1986. *A Guide to Old English*. Oxford: Basil Blackwell.
- Miura, Akira. 1979. The influence of English on Japanese grammar. *Journal of the Association of Teachers of Japanese*, 14(1): 3-31.
- Miyagawa, Shigeru. 1989. *Structure and Case Marking in Japanese: Syntax and Semantics V.22*. New York: Academic Press.

- Mizutani, Nobuko. 1985. *Nichiei hikaku hanashi kotobano bunpoo (Comparative study of grammar of conversation between English and Japanese)*. Tokyo: Kuroshio.
- Moravcsik, Edith A. 1975. Verb Borrowing. *Wieter Linguistische Gazette* 8: 3-31.
- Moravcsik, Edith A. 1978. Universals of language contact. In J. H. Greenberg (ed.), *Universals of Human Language*. Stanford: Stanford University Press. 94-122.
- Montrul, Silvina. 2004. Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and Cognition* 7(2): 125-143.
- Montrul, Silvina. 2008. *Incomplete Acquisition in Bilingualism. Re-examining the Age Factor*. John Benjamins: Amsterdam.
- Mougeon, Raymond, Nadashi, Terry & Rehner, Katherine. 2005. Contact-induced linguistic innovations on the continuum of language use: The case of French in Ontario. *Bilingualism: Language and Cognition* 8 (2): 99-115.
- Murasugi, Keiko & Sugisaki, Koji. 2008. The acquisition of Japanese syntax. In S. Miyagawa & M. Saito (eds.), *The Oxford Handbook of Japanese Linguistics*. Oxford: Oxford University Press. 250-286.
- Muysken, Pieter. 2000. *Bilingual Speech: A Typology of Code-Mixing*. Cambridge: Cambridge University Press.
- Myers-Scotton, Carol. 1993. *Duelling Languages: Grammatical Structure in Codeswitching*. Oxford: Clarendon Press.
- Myers-Scotton, Carol. 2002. *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford: Oxford University Press.
- Myers-Scotton, Carol. 2006. *Multiple voices: An introduction to bilingualism*. Malden: Blackwell Publishers.
- Nagy, Naomi. 2015. A sociolinguistic view of null subjects and VOT in Toronto heritage languages. *Lingua* 164: 309-327.
- Nagy, Naomi, G., Aghdasi, Nina, Denis, Derek & Motut, Alexandra. 2011. Null Subjects in Heritage Languages: Contact Effects in a Cross-linguistic Context. *University of Pennsylvania Working Papers in Linguistics* 17(2): 135-144.
- Namba, Kazuhiko. 2008. *English-Japanese bilingual children's code-switching: A structural approach with emphasis on formulaic language*. Doctoral dissertation, Cardiff University.
- Nariyama, Shigeko. 2000. *Referent identification for ellipted arguments in Japanese*. Doctoral dissertation. University of Melbourne, Australia.

- Nariyama, Shigeko. 2003. *Ellipsis and Reference Tracking in Japanese*. Amsterdam: John Benjamins.
- Nariyama, Shigeko. 2004. Subject ellipsis in English. *Journal of Pragmatics* 36: 237-264.
- National language research institute of Japan (Kokuritu Kokugo Kenkyusho). 1955. *Danwano jittai (Study of discourse)*, Tokyo.
- Nishimura, Miwa. 1985. *Intrasentential Code-switching in Japanese and English*. Doctoral dissertation, University of Pennsylvania.
- Nishimura, Miwa. 1986. Intrasentential code-switching: the case of language assignment. In Vaid, Jyotsna (ed.), *Language processing in Bilinguals: Psycholinguistics and Neuropsychological Perspectives*. Hillsdale, NJ: Lawrence Erlbaum Associates. 124-143.
- Nishimura, Miwa. 1995. Varietal conditioning in Japanese/English code-switching. *Language Sciences* 17(2): 123-145.
- Nishimura, Miwa. 1997. *Japanese/English Code-Switching: Syntax and Pragmatics*. New York: Peter Lang.
- Nyman, Kaisa. 2012. *English influence on Japanese popular music: A case study of language contact*. Master's Thesis. University of Eastern Finland.
- Ogawa, Akio. 2012. Case in topic-prominent language: Pragmatic and syntactic functions of cases in Japanese. In A. Malchukov & A. Spencer, *The Oxford Handbook of Case*. Oxford: Oxford University Press. 779-788.
- Okamoto, Shigeko. 1985. *Ellipsis in Japanese discourse*. Doctoral dissertation, University of California, Berkeley.
- Otheguy, Ricardo & Zentella, Ana C. 2012. *Spanish in New York: Language Contact, Dialectal Leveling, and Structural Continuity*. Oxford: Oxford University.
- Otheguy, Ricardo., Zentella, Ana C. & Livert, David. 2007. Language and dialect contact in Spanish in New York: Toward the formation of a speech community. *Language* 83(4): 770-802.
- Park, Hyeon-Sook. 2006. Structural characteristics of proper nouns in Korean-Swedish discourse. *International Journal of Bilingualism* 10(1): 17-36.
- Pavlenko, Aneta. 2004. L2 influence and L1 attrition in adult bilingualism. In M. S. Schmid, B. Köpcke, M. Keijzer, & L. Weilemar (eds.), *First Language Attrition: Interdisciplinary Perspectives on Methodological Issues*. Amsterdam/Philadelphia: John Benjamins 7-59.

- Perlmutter, David. 1971. *Deep and Surface Constraints in Generative Grammar*. New York: Holt, Rinehart and Winston.
- Polinsky, Maria. 1995. Cross-linguistic parallels in language loss. *Southwest Journal of Linguistics* 14 (1-2): 87-123.
- Polinsky, Maria. 1997. American Russian: Language loss meets language acquisition. In W. Brown, E. Dornisch, N. Kondrashova and D. Zec (eds.), *Formal Approaches to Slavic Linguistics*. Ann Arbor: Michigan Slavic Publications. 370-407.
- Polinsky, Maria. 2006. Incomplete acquisition: American Russian. *Journal of Slavic Linguistics* 14: 161–219.
- Polinsky, Maria & Kagan, Olga. 2007. Heritage languages: In the ‘Wild’ and in the classroom. *Language and Linguistic Compass* 1(5): 368-395.
- Poplack, Shana. 1980. “Sometimes I’ll start a sentence in Spanish y termino en español”: toward a typology of code-codeswitching. *Linguistics*, 7(8): 581-618.
- Poplack, Shana. 1993. Variation theory and language contact. In D. Preston (ed.), *American dialect research: An anthology celebrating the 100th anniversary of the American Dialect Society*. Amsterdam: Benjamins. 251-286.
- Poplack, Shana & Dion, Nathalie. 2012. Myths and facts about loanword development. *Language Variation and Change* 24: 279-315.
- Poplack, Shana & Levey, Stephen. 2010. Contact-induced grammatical change. In P. Auer & E. J. Schmidt (eds.), *Language and Space – An International Handbook of Linguistic Variation: Volume 1 – Theories and Methods*. Berlin: Mouton de Gruyter. 391-419.
- Poplack, Shana & Meechan, Marjory. 1995. Patterns of language mixture: Nominal structure in Wolof-French and Fongbe-French bilingual discourse. In L. Milroy & P. Muysken (eds.) *One Speaker, Two Languages: Cross-Disciplinary Perspectives on Codeswitching*. Cambridge: Cambridge University Press. 199-232.
- Poplack, Shana & Meechan, Marjory. 1998. How languages fit together in codemixing. *International Journal of Bilingualism* 2(2): 127-138.
- Poplack, Shana & Tagliamonte, Sali. 2001. *African American English in the Diaspora*. Oxford and Malden, MA: Blackwell.
- Poplack, Shana Sankoff, David & Miller, Christopher. 1988. The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26(1): 47-104.

- Poplack, Shana, Wheeler, Susan & Westwood, Anneli. 1989. Distinguishing language contact phenomena: evidence from Finnish-English bilingualism. *World Englishes* 8: 389-406.
- Poplack, Shana, Zentz, Lauren & Dion, Nathalie. 2012a. Phrase-final prepositions in Quebec French: An empirical study of contact, code-switching and resistance to convergence. *Bilingualism: Language and Cognition* 15(2): 203-225.
- Poplack, Shana, Zentz, Lauren & Dion, Nathalie. 2012b. What counts as (contact-induced) change. *Bilingualism: Language and Cognition* 1: 247–254.
- Preston, Dennis R. 1986. The case of American Polish. In D. Kastovsky & A. Szwedek (eds.), *Linguistics across Historical and Geographical Boundaries. In Honour of Jacek Fisiak on the Occasion of his Fiftieth Birthday*. Vol. 2. Berlin: Mouton de Gruyter. 1015-23.
- Raña Risso, R. 2010. Subject pronoun placement as evidence of contact and leveling in Spanish in New York. *International Journal of the Sociology of Language* 203: 101–114.
- Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.
- Roberts, Ian & Holmberg, Anders. 2010. Introduction: parameters in minimalist theory. In T. Biberauer, A. Holmberg, I. Roberts & M. Sheehan (eds.), *Parametric Variation: Null Subjects in Minimalist Theory*. New York: Cambridge University Press. 1-57.
- Romaine, Suzanne. 1986. The syntax and semantics of the code-mixed compound verb in Panjabi-English bilingual discourse. In D. Tannen & J. Alatis (eds.), *Language and Linguistics: The Interdependence of Theory, Data and Application*. Washington, DC: Georgetown University Press. 35-49.
- Saito, Mamoru. 1983. Case and Government in Japanese. *Proceedings of the West Coast Conference on Formal Linguistics* 2: 247-259.
- Saito, Mamoru. 1985. *Some Asymmetries in Japanese and Their Theoretical Implications*. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, USA.
- Samar, Reza G. & Meechan, Marjory. 1998. The null theory of codeswitching versus the nonce borrowing hypothesis: Testing the fit in Persian-English bilingual discourse. *International Journal of Bilingualism* 2(2): 203-220.
- Sánchez, Liliana. 2004. Functional convergence in the tense, evidentiality and aspectual systems of Quechua Spanish bilinguals. *Bilingualism: Language and Cognition* 7(2): 147-162.
- Sankoff, David. 1988. Variable rules. In U. Ammon, N. Dittmar & K. J. Mattheier (eds.), *Sociolinguistics: An International Handbook of the Science of Language and Society Vol. 2*. Berlin and New York: Walter de Gruyter. 984-997.

- Sankoff, David & Poplack, Shana. 1981. A formal grammar for code-switching. *Papers in Linguistics* 14(1): 3-46.
- Sankoff, David, Poplack, Shana, & Vanniarajan, Swathi. 1990. The case of the nonce loan in Tamil. *Language Variation and Change* 2: 71-101.
- Sankoff, David, Tagliamonte, Sali, & Eric Smith. 2005. Goldvarb X: A multivariate analysis application for Macintosh and Windows. Toronto, Canada: Department of Linguistics, University of Toronto. <http://individual.utoronto.ca/tagliamonte/goldvarb.html>
- Sankoff, Gillian & Labov, William. 1985. Variation theory. In *Papers from New Ways of Analyzing Variation (N WAV) 14*, Georgetown University, Washington D.C.
- Schilling-Estes, Natalie. 2007. Sociolinguistic fieldwork. In R. Bayley & C. Lucas (eds.), *Sociolinguistic Variation: Theory, Methods and Application*. Cambridge: Cambridge University Press. 165-189.
- Schmid, Monika S. 2002. *First Language Attrition, Use and Maintenance: The Case of German Jews in Anglophone Countries*. Amsterdam: John Benjamins.
- Schmid, Monika S., Köpcke, Barbara, Keijzer, Merel & Weilemar, Lin. 2004. *First Language Attrition: Interdisciplinary Perspectives on Methodological Issues*. Amsterdam: Johan Benjamins.
- Scott, Katherine J. 2010. *The Relevance of Referring Expressions: The Case of Diary Drop in English*. Doctoral dissertation. University College, London.
- Seliger, Herbert W. & Vago, Robert M. 1991. *First Language Attrition*. Cambridge: Cambridge University Press.
- Shibamoto, Janet S. 1980. *Language Use and Linguistic Theory: Sex-related Variation in Japanese Syntax*. Doctoral dissertation. University of California at Davis.
- Shibamoto, Janet S. 1983. Subject ellipsis and topic in Japanese. *Papers in Linguistics* 16(1-2): 233-265.
- Shibamoto, Janet S. 1985. *Japanese Women's Language*. New York: Academic.
- Shibatani, Masayoshi. 1990. *The Language of Japan*. Cambridge: Cambridge University Press.
- Silva-Corvalán, Carmen. 1982. Subject expression and placement in Mexican-American Spanish. In J. Amaste & L. Elias-Olivares (eds.), *Spanish in the United States: Sociolinguistic aspects*. New York: Cambridge University Press. 93-120.
- Silva-Corvalán, Carmen. 1994. *Language Contact and Change: Spanish in Los Angeles*. New York: Oxford University Press.

- Silva-Corvalán, Carmen. 2001. *Sociolingüística y Pragmática del Español*. Washington, D.C.: Georgetown University Press.
- Smith, Donald L. 1978. Mirror Images in Japanese and English. *Language* 54 (1): 78-122.
- Sorace, Antonella. 2004. Native language attrition and developmental instability at the syntax-discourse interface: data, interpretations and methods. *Bilingualism: Language and Cognition* 7: 143-5.
- Sorace, Antonella. 2005. Syntactic optionality at interfaces. In L. Cornips & K. Corrigan (eds.), *Syntax and variation: Reconciling the biological and the social*. Amsterdam: John Benjamins. 46-111.
- Sorace, Antonella. 2011. Cognitive advantages of bilingualism: is there a “bilingual paradox”? In P. Valore (ed.), *Multilingualism. Language, Power, and Knowledge*. Edistudi: Pissa. 335-358.
- Sorace, Antonella & Serratrice, Ludovica. 2009. Internal and external interfaces in bilingual language development: beyond structural overlap. *International Journal of Bilingualism* 13(2): 195-210.
- Stanlaw, James. 1987. Japanese and English: Borrowing and contact. *World Englishes* 6(2): 93-109.
- Stanlaw, James. 2004. *Japanese English: Language and culture contact*. Hong Kong: Hong Kong University Press.
- Takano, Shoji. 1998. A quantitative study of gender difference in the ellipsis of the Japanese postpositional particles *-wa* and *-ga*: Gender composition as a constraint on variability. *Language Variation and Change* 10: 289-323.
- Tagliamonte, Sali A. 2006. *Analysing Sociolinguistic Variation*. New York: Cambridge University Press.
- Tagliamonte, Sali A. 2012. *Variationist Sociolinguistics: Change, Observation, Interpretation*. Oxford: Blackwell.
- Tamori, Ikuhiro. 1977. NP and particle deletion in Japanese discourse. *Southern California Occasional Papers in Linguistics* 5: 243-263.
- Taraldsen, Knut T. 1978. *On the Nominative Island Condition, Vacuous Application and the That-trace Filter*. Ms., MIT, Cambridge, Massachusetts.
- Thomason, Sarah G. 2001. *Language Contact: An Introduction*. Edinburgh: Edinburgh University Press.

- Thomason, Sarah G. 2003. Contact as a source of language change. In B. D. Joseph & R. D. Janda (eds.), *The Handbook of Historical Linguistics*. Oxford: Blackwell. 687-712.
- Thomason, Sarah G. & Kaufman, Terrence. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Los Angeles: University of California Press.
- Tomioka, Satoshi 2003. The semantics of Japanese null pronouns and its cross-linguistic implications. In K. Schwabe & S. Winkler (eds.), *The Interfaces: Deriving and Interpreting Omitted Structures*. Amsterdam: John Benjamins. 321- 339.
- Toribio, Jacqueline A. 2004. Convergence as an optimization strategy in bilingual speech: Evidence from code-switching. *Bilingualism: Language and Cognition* 7(2): 165- 173.
- Torres Cacoullous, Rena & Aaron, Jessie E. 2003. Bare English-origin nouns in Spanish: Rates, constraints, and discourse functions. *Language Variation and Change* 15(3): 287-326.
- Torres Cacoullous, Rena & Travis, Catherine E. 2010. Variable yo expression in New Mexico: English influence? In S. Rivera-Mills & D. J. Villa (eds), *Spanish of the U.S. Southwest: A Language in Transition*. Madrid: Iberoamericana. 185-206.
- Tóth, Gergely. 2007. *Linguistic Interference and First-Language Attrition: German and Hungarian in the San Francisco Bay Area*. New York: Peter Lang.
- Travis, Catherine E. & Torres Cacoullous, Rena. 2015. Two languages, one effect: Structural priming in spontaneous code-switching. *Bilingualism: Language and Cognition* 18. Advance online publication. doi: 10.1017/S1366728914000406
- Treffers-Daller, Jeanine. 2005. Evidence for insertional codemixing: Mixed compounds and French nominal groups in Brussels Dutch. *International Journal of Bilingualism* 9(3-4): 477-508.
- Tsao, Feng-fu. 1977. *A Functional Study of Topic in Chinese: The First Step toward Discourse Analysis*. Doctoral dissertation, University of Southern California.
- Tsujimura, Natsuko. 1990. Ergativity of nouns and case assignment. *Linguistic Inquiry* 21: 277-287.
- Tsujimura, Natsuko. 1996. *An Introduction to Japanese Linguistics*. Cambridge: Blackwell.
- Tsutsui, Michio. 1984. *Particle Ellipses in Japanese*. Doctoral dissertation, University of Illinois.
- Turpin, Danielle. 1998. 'Le français c'est le last frontier': The status of English-origin nouns in Acadian French. *International Journal of Bilingualism* 2(2): 221-233.
- Van Coetsem, Frans. 2000. *A General and Unified Theory of the Transmission Process in Language Contact*. Hridelberg: Winter.

- Weinreich, Uriel. 1953. *Languages in Contact: Findings and Problems*. New York: Publications of the Linguistic Circle of New York.
- Weinreich, Uriel, Labov, William Labov, & Herzog, Marvin. 1968. Empirical foundations for a theory of language change. In W. P. Lehmann & Y. Malkiel (eds.), *Directions for Historical Linguistics*. Austin, TX: University of Texas Press. 95-189.
- Weir, Andrew. 2009. Subject pronoun drop in informal English. Richard M. Hogg Prize winning essay. <http://people.umass.edu/aweir/>
- Winford, Donald. 2003. *An Introduction to Contact Linguistics*. Malden: Blackwell.
- Winford, Donald. 2009. On the unity of contact phenomena and their underlying mechanisms. In L. Isurin, D. Winford & K. de Bot (eds.), *Multidisciplinary Approaches to Codeswitching*. Amsterdam: John Benjamins. 279-305.
- Yoshizumi, Yukiko. 2012. *The Status of Lone English-origin Nouns in Otherwise Japanese Discourse: A Comparative Variationist Approach to Borrowing*. Comprehensive Exam Paper. University of Ottawa.
- Zentella, Ana C. 1997. *Growing Up Bilingual*. Malden, MA: Blackwell.