ADAPTATION AND TRANSMISSION IN JAPANESE LOANWORD PHONOLOGY

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Clifford James Crawford

August 2009

ADAPTATION AND TRANSMISSION IN JAPANESE LOANWORD PHONOLOGY

Clifford James Crawford, Ph.D.

Cornell University 2009

This dissertation focuses on two issues in loanword phonology: first, how are loanwords represented in the lexicons of L1 speakers; and second, how do loanword adaptation patterns become established in a speech community over time. In examining these questions, I propose a theoretical framework for loanword borrowing in which the nativization of loanwords can take place during either of two processes: the initial adaptation of a new loanword by a borrower, or the transmission of loanwords from speaker to speaker within a social network.

Using data from Arakawa's (1977) dictionary of English loanwords in Japanese, I show that historical adaptation patterns are affected not only by native phonological constraints, but also by type frequency and phonological similarity to other loanwords. Non-native phonotactic sequences are more likely to be preserved in more common phonological environments than in less common environments. I present an OT analysis of the historical variation in adaptation conventions, showing that the set of possible adaptations for a loanword can be obtained by reranking loanword-specific faithfulness constraints against a fixed ranking of native markedness constraints. However, this analysis cannot explain the origin of frequency and phonological neighborhood effects on adaptations. To account for these effects, I argue for a model of the lexicon in which lexical entries are organized on the basis of phonological similarity. I then develop a connectionist model of loanword adaptation to show how native constraints interact with type frequency in the adaptation of new loanwords.

As for the establishment of adaptation patterns in a speech community, I argue that the transmission of loanwords from one speaker to another plays an important role. Based on recent research on the structure of social networks, I develop an agent-based model of loanword transmission within a network of speakers, showing that in general transmission causes a nonlinear amplification of the effects of nativization at the level of the individual speaker. I apply this model to the attested rates of nativization of various non-native patterns in Japanese loanwords, showing that Japanese speakers before 1890 tended to palatalize coronal stops before /i/ in loanwords at a much greater rate than they nativized other non-native patterns. After 1890, the attested rate of coronal palatalization in new loanwords drops to roughly the same rate as other nativizations. This data suggests that before 1890, coronal palatalization was a categorical process for many Japanese speakers, whereas after 1890, perhaps because of increased English-language education at this time, /ti/ and /di/ became allowable sequences in loanwords and palatalization became a low-level gradient process occurring during loanword transmission.

BIOGRAPHICAL SKETCH

Clifford James Crawford was born in Syracuse, New York on July 26, 1978. He graduated from Corcoran High School in June of 1996 and attended Cornell University as an undergraduate. He graduated from Cornell in May of 2000 with a B.A. in Linguistics (concentration in Computer Science). He stayed in the Linguistics program at Cornell, receiving an M.A. in March of 2004. After spending the summer of 2004 studying in a Japanese immersion program at Middlebury College, he returned to Cornell for his dissertation research, and successfully defended his dissertation in November of 2008. He is currently working as a research scientist at Textwise, LLC in Syracuse, New York.

for Keith

and Anabel

ACKNOWLEDGEMENTS

There are many, many people I would like to thank who have helped me over the course of my graduate career:[1]

- My advisor, John Whitman, for agreeing to take me on as a student even though I was making the insane proposal to switch subfields and write about a language I hadn't yet studied, and for teaching me just about everything I know about the structure and history of the Japanese language. Without him this project would never even have been started, or seen through to its completion.
- The other members of my committee, Abby Cohn, Morten Christiansen, and Michael Wagner, for giving me a variety of perspectives on phonology and cognitive science, making my dissertation an even stronger piece of work.
- Barbara Lust, for providing me with employment in the Cornell Language Acquisition Lab for much of the time I've been here, and the many members of the lab that have come and gone over the years, including María Blume, Whitney Postman, and Katherine White.
- The many other professors at Cornell that I've had an opportunity to learn from in my time here. Thank you in particular to Wayne Harbert, for getting me interested in language and linguistics in the first place as an undergrad; to Chris Collins, for serving as my advisor for my first few years in grad school and for showing me how to think about syntax and linguistic theory in general; to Shimon Edelman, for arousing my interest in cognitive science and presenting a very different view of how language might work; and to Michael Spivey, for graciously doing demos in his eye-tracking lab for my freshman writing students every semester I taught that course.

[1] This research was supported in part by NSF grant #0741666, "Automaton Theories of Human Sentence Comprehension".

- The members of my dissertation writing group, Alexa Yesukevich, Baran Han, Heather MacLachlan, and Matthew Hoffberg, for keeping me motivated and on track throughout the lead-up to the defense, and the many revisions afterwards.
- My friends and fellow grad students in the Cornell Linguistics department, for enduring my (increasingly frequent) rants in the Pain Cave over the years: Edith Aldridge, Gabe Arana, Julie Balazs, Christina Bjorndahl, Marisa Boston, Johanna Brugman, Ed Cormany, Sarah Courtney, Serena Crivellaro, Alison Fisher (you can have your chair back now), Tova Friedman, Teresa Galloway, Effi Georgala, Masa Gibson, Kyle Grove, Rachel Hastings, Dan Kaufman, Steven Ikier, Satoshi Ito, Ellert Þór Jóhannsson, Andrew Joseph, Brandi Morgan, Peggy Renwick, Kazuha Watanabe, and many others who I don't have space to list here. Thank you especially to Adam Cooper for being a real mensch in all of the years that I've known him, and to Becky Butler Thompson for not killing me (yet).
- The many other people I've befriended during my time at Cornell, for reminding me that there is a world outside of grad school, including Anna, Lyn Arcega, Mark and Aja Bogdanoff, Manuel Calimlim, Charlie, Hope, Kris Gilbert, Mariko Molander, Nadine Latief, Natalie Liu, Molly, Yohei Nakayama, Robert Schaufelberger, RuDy! Ma, Yuko Torikata, 月子 Tsukiko, Anne "餡こ" Warlaumont, Wendy, and Ryan Yang.
- My mother Walsi, my sisters Jessica and Meagan, my brother John, my brother-in-law Richard Peck, and the rest of my family, for their patience and support throughout this long process. I only wish that my grandparents, James and Shirley Felker, and my stepfather, Patrick King, could have lived to see the end of this. I'm sure they would have been very proud.

TABLE OF CONTENTS

Biographical Sketch iii
Dedication iv
Acknowledgements v
Table of Contents vii
List of Tables ix
List of Figures xii

1 Introduction 1
1.1 Loanword phonology 1
1.2 Lexical strata in Japanese 3
1.2.1 *Rendaku* and Lyman's Law 7
1.2.2 *NT and *P 8
1.3 Japanese loanword adaptation patterns 12
1.4 Outline for the remainder of the dissertation 18
1.5 A note on terminology 21

2 Adaptation and transmission of loanwords 23
2.1 The borrowing process 23
2.2 Nativization during transmission 28
2.3 Changes in adaptation patterns in Japanese loanwords 35
2.3.1 Loanword-specific faithfulness constraints 35
2.3.2 Data collection 38
2.3.3 Palatalization of velars before /æ/ 41
2.3.4 Gemination of word-final voiced obstruents 57
2.3.5 Coronals and palatals before front vowels 69
2.4 General observations 80
2.5 Conclusion 83

3 Generative models of loanword adaptation and representation 85
3.1 The input to loanword adaptation 86
3.2 Nativization during the SP and CS stages 94
3.3 Lexical stratification 97
3.3.1 The Core-Periphery model of lexical stratification 98
3.3.2 The Cophonology model 110
3.4 The representation of phonological neighborhoods 113
3.5 Conclusion 117

4 A connectionist model of loanword adaptation 119
4.1 Connectionist models of cognition 120
4.2 Constraint-based theories of phonology and lexical processing 127
4.3 A connectionist framework for lexical representation in Japanese 140
4.3.1 Feature representation 141

4.3.2 Stratum classification . . . 145
4.4 Stratum classification network . . . 150
4.4.1 The similarity structure of the Japanese lexicon . . . 153
4.4.2 LDA classifier . . . 160
4.4.3 Feedforward network classifier . . . 163
4.4.4 Discussion . . . 167
4.5 Loanword adaptation network . . . 169
4.5.1 Methods . . . 170
4.5.2 Results . . . 178
4.5.3 Discussion . . . 183
4.6 Conclusion . . . 189

5 Loanword transmission in social networks **191**
5.1 Characterizing social structure . . . 192
5.1.1 Social network analysis in sociolinguistics . . . 194
5.1.2 The large-scale structure of social networks . . . 199
5.2 A formal framework for modeling loanword borrowing . . . 203
5.2.1 Transmission in line graphs . . . 205
5.2.2 Transmission in random graphs . . . 214
5.3 Modeling Japanese adaptation patterns . . . 220
5.4 Conclusion . . . 230

6 Conclusion **231**
6.1 Phonological grammars and connectionist models of lexical processing 234
6.2 Agent-based simulations of language contact and language change . . . 237

Bibliography **240**

LIST OF TABLES

1.1 Phonological and phonetic values for *hiragana* and *katakana* characters 5
1.2 Examples of Yamato, Sino-Japanese, and Foreign words 6
1.3 Examples of *rendaku* 7
1.4 Past tense alternations in Yamato 9
1.5 Counter words 10
1.6 The /maQ−/ adjectival prefix 10
1.7 Examples of *rendaku* with h → b 11
1.8 Vowel systems of Japanese and American English 13
1.9 Consonant inventory of Japanese 14
1.10 Adaptation of *bike*→[baikɯ] 14
1.11 Possible voiceless coronal-vowel and palatal-vowel sequences in Japanese 15
1.12 Palatalization in Japanese verb alternations 15
1.13 Adaptation of *dilemma*→[ʤireɴma] 16

2.1 Adaptation of *team*→[ʧiːmɯ], *[tiːmɯ] 29
2.2 Adaptation of *best*→[besɯto], showing loanword-specific epenthesis repairs 37
2.3 Deletion repairs in verb morphology 37
2.4 Adaptation patterns for KÆ words, by date of first attestation and cognate status 47
2.5 Adaptation of /kæ/→[ke], with IDENT-SB(back) ≫ IDENT-SB(low) . 48
2.6 Adaptation of /kʲæ/→[kʲa] 49
2.7 Adaptation of /kʲæ/→[ke], with *Cʲa ≫ IDENT-SB(back) 50
2.8 Incorrect adaptation of *cat*→[katto] instead of [kʲatto] 51
2.9 Palatalization in loanword doublets 53
2.10 Adaptation of *café*→[kaɸe] 56
2.11 Adaptation of *pet*→[petto] 58
2.12 Adaptation of *date*→[deːto], *[deːtto] 58
2.13 Gemination and epenthesis alternations in Sino-Japanese compounding 59
2.14 Gemination and nasal-obstruent alternations in Mimetic roots 60
2.15 Gemination and nasal-obstruent alternations in past tense forms 60
2.16 Adaptation of *bed*→[betto], with *DD ≫ IDENT-SB(voi) 61
2.17 Adaptation patterns for voiced geminates, by date of first attestation and final consonant 64
2.18 Adaptation of *bed*→[beddo], with {ALIGN-SB, IDENT-SB(voi)} ≫ *DD 65
2.19 Adaptation of *pub*→[pabɯ], with {*DD, IDENT-SB(voi)} ≫ ALIGN-SB 66
2.20 Adaptation of *pub*→[pabɯ], with *BB ≫ IDENT-SB(voi) ≫ ALIGN-SB ≫ *DD 67
2.21 Adaptation of *team*→[ʧiːmɯ], *[tiːmɯ] 69
2.22 Adaptation of *shepherd*→[sepaːdo], *[ʃepaːdo] 70
2.23 *TI violations and attestations given in Arakawa (1977) 71

2.24 Type frequencies of phonological environments for TI-containing loanwords . . . 77
2.25 The adaptation pattern TI→TI, with {IDENT-SB(cont), IDENT-SB(high)} ≫ *TI . . . 78
2.26 The adaptation pattern TI→ČI, with {*TI, IDENT-SB(high)} ≫ IDENT-SB(cont) . . . 78
2.27 The adaptation pattern TI→TE, with {*TI, IDENT-SB(cont)} ≫ IDENT-SB(high) . . . 79
2.28 Adaptation of *director*→[direkɯtaː], with *TI ≫ IDENT-SB(cont) ≫ *DI . . . 79

3.1 Deletion and epenthesis in Japanese loanword doublets . . . 93
3.2 Pattern of violations for *P and *NT . . . 99
3.3 Pattern of violations for *P and *NT, including Mimetic . . . 101
3.4 Phonological constraints distinguishing lexical strata . . . 106
3.5 Yamato *NT violations . . . 107
3.6 Adaptation patterns of English monosyllables . . . 109
3.7 Examples of truncated loanwords . . . 109
3.8 Constraint rankings for Japanese lexical strata . . . 111
3.9 Constraint rankings for assimilated and unassimilated loanwords . . . 111

4.1 Constraint-based models of phonology and the lexicon . . . 132
4.2 Harmonic Grammar analysis of geminate devoicing in Japanese loanwords . . . 133
4.3 The major class features [son] and [cons] . . . 141
4.4 The features [cont_1] and [cont_2] . . . 143
4.5 Articulator features, [ant], and [high] . . . 143
4.6 Feature encoding for Japanese segments . . . 146
4.7 Example vector representations for training items . . . 147
4.8 Lexical strata derived from JMDICT . . . 150
4.9 Distribution of syllable types for four-mora words in the Yamato, Sino-Japanese, and Foreign strata . . . 152
4.10 Stratum classification training and test data . . . 160
4.11 Confusion matrix for results from LDA classifier . . . 161
4.12 Confusion matrix for results from stratum classification network . . . 164
4.13 Mean shortest distance from Foreign test items to Yamato, Sino-Japanese, and Foreign training items . . . 166
4.14 Mora alignment on input layer of adaptation network . . . 171
4.15 Training and test data for adaptation networks . . . 177
4.16 Adaptation network performance on training and testing items . . . 178
4.17 Test words for adaptation networks . . . 179
4.18 Adaptation patterns of voiced geminates produced by Network 1 . . . 182
4.19 Adaptation patterns of voiced geminates produced by Network 2 . . . 184
4.20 Adaptation patterns of voiced geminates produced by Network 3 . . . 185

4.21 Adaptation patterns of voiced geminates produced by Network 4 186
4.22 Most common adaptation patterns produced by networks for each test loanword 187

5.1 Typical values of graph parameters for random and social networks . . 199
5.2 Mean values of $n_1 \ldots n_5$ in $G(N, p)$ random graphs 217

LIST OF FIGURES

2.1 The game of Telephone played with loanwords 30
2.2 Adaptation patterns for velar stops before /æ/, by year of first attestation 43
2.3 Adaptation patterns for velar stops before /æ/, by cognate status 44
2.4 KYA and KA variants of a loanword spreading in a simplified social network. 54
2.5 Adaptation patterns of word-final voiced obstruents, by date of first attestation 62
2.6 Adaptation patterns of word-final voiced obstruents, by final consonant 63
2.7 Adaptation patterns over time of *TI-violating loanwords 72
2.8 Adaptation patterns over time of source words containing /t/ 75
2.9 Adaptation patterns over time of source words containing /d/ 75
2.10 Adaptation patterns by phonological neighborhood for TI sequences in loanwords first attested 1870–1899 76
2.11 Adaptation patterns by phonological neighborhood for TI sequences in loanwords first attested 1900–1929. 76
2.12 Estimated nativization rates for coronals before /i/, palatals before /e/, voiced geminates, and contrastive /ɸ/ 82
2.13 Estimated nativization rates per 20-year period for coronals before /i/, palatals before /e/, voiced geminates, and contrastive /ɸ/ 82

3.1 The SP and CS stages in loanword adaptation 85
3.2 Sources of variation in loanword adaptation 92
3.3 The Core-Periphery model of the lexicon 100
3.4 The Core-Periphery model of the lexicon (with Mimetic added) 101
3.5 The minimum length constraint on Mimetic roots, with Sino-Japanese at the core of the lexicon 104
3.6 Domains of phonological constraints 106
3.7 Grammar lattice for Japanese lexical strata 112

4.1 The logistic and tanh functions 121
4.2 A feedforward connectionist network 122
4.3 An attractor network 123
4.4 The Triangle Model of lexical processing 129
4.5 Features used in network simulations 142
4.6 Slot-based representation for Japanese phonology 144
4.7 Schematic representations of possible lexicon organization patterns . . 154
4.8 MDS analysis of feature vector representations for random sample of lexical items from `JMDICT` 156
4.9 Scatterplot of MDS axes 1 and 2 from Figure 4.8 157
4.10 Scatterplot of MDS axes 1 and 3 from Figure 4.8 158
4.11 Network architecture for the stratum classification network 163
4.12 Principal component analysis of hidden unit activations in classification network 165

4.13 Architecture of the adaptation network 172
4.14 Frequencies of words containing geminate obstruents in training set for Network 2 175
4.15 Frequencies of words containing geminate obstruents in training set for Network 4 176

5.1 A hypothetical social network 193
5.2 Log-linear plot of cumulative degree distributions of random and scale-free graphs 202
5.3 Loanword transmission in a social network 204
5.4 A line graph with *N* speakers 205
5.5 Predicted values of q_f for various values of p_f in a line network 211
5.6 Results of computer simulations of loanword transmission in line graphs 213
5.7 A tree of *N* nodes 215
5.8 Splitting a graph into a set of trees rooted at the initial borrowers 217
5.9 Mean values of $n_1 \dots n_5$ in random graphs 218
5.10 Multiple paths between a node and the initial borrower in a graph . . . 219
5.11 Results of simulations of loanword transmission in random graphs . . . 221
5.12 Results of simulations of loanword transmission in small-world graphs 222
5.13 Ratios of F tokens to total attested tokens per 20-year period for coronals before /i/, palatals before /e/, voiced geminates, and contrastive /ɸ/ 223
5.14 Finding p_f values using polynomial regression on the results from Figure 5.12 226
5.15 Estimated values of p_f for the historical nativization data in a small-world graph 227

CHAPTER 1

INTRODUCTION

1.1 Loanword phonology

The borrowing of words from one language into another has long been a topic of interest among linguists. One reason for this is that the phonological nativizations seen in loanwords are a useful source of extragrammatical evidence (along with language games, poetics, speech rate and style variation, and so forth) about how the phonological system of the borrowing language operates. For example, the borrowing into English of Japanese words containing /a/, such as 空手 /kaɾate/→[kəˈɹɑti], 山葵 /wasabi/→[wəˈsɑbi], and やくざ /yakɯza/→[ˈyɑkuzə], shows that in general Japanese /a/ is adapted as [ɑ] when it occurs in a stressed syllable in the resulting English loanword,[1] but is instead [ə] when it occurs in an unstressed syllable. This suggests that there is a restriction in English on /ɑ/ surfacing in unstressed open syllables. This putative constraint can then be corroborated using language-internal evidence such as static distribution patterns in the lexicon and phonological alternations. Many researchers have thus argued that loanword nativization patterns can provide additional insight, beyond grammar-internal evidence, into the phonological competence of the speakers of the borrowing language (Kenstowicz & Suchato 2006).

Borrowings also have important implications for our understanding of the structure of the mental lexicon. Loanword phenomena often make it difficult to characterize the phonology of a language as a single unified system operating indiscriminately on all of the items in the lexicon. The analysis by Fries & Pike (1949) of post-nasal voicing in

[1]Though there are some cases, like 侍 /samɯɾai/→[ˈsæməɹɑɪ] or the British pronunciation of 酒 /sake/ as [ˈsæki], where /a/ is adapted as [æ] instead.

Mazatec provides a clear example of this. In native Mazatec words, voiceless stops are always voiced when they occur after a nasal, suggesting an analysis in which voiced and voiceless stops (e.g. [t] and [d]) are allophones of a single phoneme (/t/). However, there are a very small number of loans from Spanish, such as *siento* 'one hundred', in which a voiceless stop occurs after a nasal. Fries & Pike argue that there must be two "coexistent phonemic systems" in Mazatec: the native system, in which stops are obligatorily voiced after nasals, and the loanword system, in which a voicing contrast is maintained after nasals. Thus loanword borrowing can end up introducing new segments and phonemic contrasts that did not previously exist in the recipient language. In some cases these can even spread to the native stratum of the vocabulary, as with the allophonic voicing alternations in fricatives in Middle English, which did not become contrastive until large numbers of loanwords were borrowed from Norman French and Old Norse.

A third topic of interest, one which has not been well studied in the generative literature on loanwords, is the diachronic establishment of adaptation patterns in a speech community. At the beginning stages of contact between L1 (the recipient language) and L2 (the donor language), there is often a great amount of variability in attested adaptation patterns. Yet in long-term contact situations, this variability is resolved over time into a coherent set of conventions for borrowing words from L2 which all L1 speakers share (Haugen 1950). How this process of regularization arises from the interactions between L1 speakers is an important question not only for the specific case of loanword borrowing, but also for the larger question of how language change in general is implemented (cf. the "actualization" problem of Weinreich, Labov & Herzog 1968).

Thus, loanword borrowing holds much interest for linguistic theory because it sheds light on the phonological competence of individual speakers, on the structure of the mental lexicon, and on the process of conventionalization in speech communities. Japanese is a particularly interesting case to study with respect to loanword phenomena because it exemplifies all three of these aspects of borrowing. Japanese has, over its history, been subject to large-scale borrowing from two very different languages phonologically (Middle Chinese and modern-day English), resulting in a lexicon with two distinct loanword strata, each with their own unique phonological properties distinguishing them from the native and mimetic strata of the lexicon. Psycholinguistic experiments such as those conducted by Dupoux, Kakehi, Hirose, Pallier & Mehler (1999) testing the adaptation of novel loanwords have been conducted with modern-day Japanese speakers, allowing us to look at the effects of perception and phonology at the level of an individual speaker. And, borrowing from English into Japanese has been taking place over the past 150 years, and continues to this day, providing an ample source of written evidence for how the present-day conventions for borrowing words from English were originally established.

In the following section I will present a brief description of some of the phonological and morphological processes governing the different lexical strata of Japanese, leading up to a discussion of recent loanwords from English and their interest for phonological theory.

1.2 Lexical strata in Japanese

Japanese is traditionally described as having four distinct lexical strata (Martin 1952, McCawley 1968).[2] The native vocabulary which can be traced back to Old Japanese is

[2] An interesting exception is Bloch (1950), who does not consider even recent loanwords in Japanese or other languages as constituting a "separate phonemic system". He argues specifically against the analysis of Mazatec in Fries & Pike (1949) as containing two coexistent phonological systems, one which applies to native words, and the other to loanwords from Spanish. From Bloch's point of view, all members of the lexicon are valid data for phonemic analysis, regardless of their etymology or the length of time they have existed in the lexicon. Although Bloch does not state this explicitly, this principle would lead to the

known as the *Yamato* stratum (from a Japanese word meaning ‘native’). Yamato items are extremely common in natural speech and in various genres of written text (Shibatani 1999: 142–3). Basic vocabulary, like words for plants, animals, body parts, and colors, as well as most verbs, adjectives, adverbs, and grammatical words like conjunctions and case particles, are in the Yamato stratum. Yamato words are written using *hiragana* (Table 1.1), one of two moraic writing systems used in Japanese, and *kanji*, a set of characters borrowed from Chinese.

Loanwords which were borrowed from Middle Chinese in the 7th–13th centuries are known as *Sino-Japanese*. These comprise a sizeable portion of the lexicon, and are somewhat analogous in status to the Latinate vocabulary of English (Shibatani 1999: 145–6). Sino-Japanese words are made up of mono- or bisyllabic units, called *roots*. Each root corresponds to a single *kanji* character in the writing of the word. For example, the word 言語学 /geɴgogakɯ/[3] ‘linguistics’ is made up of three roots, /geɴ/, /go/, and /gakɯ/, corresponding to the three characters in the word, 言, 語, and 学. In some cases, these roots correspond to distinct morphemes. For example, 学 /gakɯ/ often occurs in names of fields of study like 言語学 /geɴgogakɯ/ ‘linguistics’, 数学 /sɯːgakɯ/ ‘mathematics’, 歴史学 /rekiʃigakɯ/ ‘history’, and so on. But not all Sino-Japanese words can be transparently derived from the meanings of their constituent roots. For instance, the semantic relationship between the roots 勉 /beɴ/ ‘exertion’ and 強 /kʲoː/ ‘strong’ and the word 勉強 /beɴkʲoː/ ‘study’ is obscure at best.

Loanwords which were borrowed from the 16th century onward make up the *For-*

---

conclusion that Mazatec makes a voicing distinction in post-nasal stops, without any qualifications; the fact that this contrast is made only in a small subset of the lexicon would be deemed irrelevant. However, even Bloch recognizes that there are significant generalizations which can be made over the portion of the Japanese lexicon not containing recent loanwords. He describes these generalizations in terms of two types of idiolects in use at the time among speakers of standard Japanese: the “conservative” dialect in which all loanwords are nativized, and the “innovating” dialect in which loanwords can contain non-native phonotactics.

[3] /ɴ/ represents a placeless nasal, which is realized phonetically in various ways, depending on its context (Akamatsu 1997). Here it assimilates in place to the following stop and is realized as [ŋ].

Table 1.1: Phonological and phonetic values for *hiragana* and *katakana* characters. Characters in parentheses are obsolete and rarely used today. Obstruent voicing is indicated by adding a *dakuten* mark to the corresponding voiceless *kana*; for example, ⟨た⟩ is the *hiragana* character for /ta/, while ⟨だ⟩ is the character for /da/. A *handakuten* mark is added to the characters in the /h/ column to denote /p/: ⟨ぱ⟩, ⟨ぴ⟩, ⟨ぷ⟩, ⟨ぺ⟩, and ⟨ぽ⟩ are the *hiragana* characters for /pa/, /pi/, /pɯ/, /pe/, and /po/.

		/w–/	/ɾ–/	/j–/	/m–/	/h–/	/n–/	/t–/	/s–/	/k–/	Ø	
hiragana		わ	ら	や	ま	は	な	た	さ	か	あ	/–a/
katakana		ワ	ラ	ヤ	マ	ハ	ナ	タ	サ	カ	ア	
		[wa]	[ɾa]	[ja]	[ma]	[ha]	[na]	[ta]	[sa]	[ka]	[a]	
hiragana		(ゐ)	り		み	ひ	に	ち	し	き	い	/–i/
katakana		(ヰ)	リ		ミ	ヒ	ニ	チ	シ	キ	イ	
		[i]	[ɾi]		[mi]	[çi]	[ni]	[cɕi]	[ɕi]	[ki]	[i]	
hiragana			る	ゆ	む	ふ	ぬ	つ	す	く	う	/–ɯ/
katakana			ル	ユ	ム	フ	ヌ	ツ	ス	ク	ウ	
			[ɾɯ]	[jɯ]	[mɯ]	[ɸɯ]	[nɯ]	[tsɯ]	[sɯ]	[kɯ]	[ɯ]	
hiragana		(ゑ)	れ		め	へ	ね	て	せ	け	え	/–e/
katakana		(ヱ)	レ		メ	ヘ	ネ	テ	セ	ケ	エ	
		[e]	[ɾe]		[me]	[he]	[ne]	[te]	[se]	[ke]	[e]	
hiragana	ん	を	ろ	よ	も	ほ	の	と	そ	こ	お	/–o/
katakana	ン	(ヲ)	ロ	ヨ	モ	ホ	ノ	ト	ソ	コ	オ	
	/ɴ/	[o]	[ɾo]	[jo]	[mo]	[ho]	[no]	[to]	[so]	[ko]	[o]	

Table 1.2: Examples of Yamato, Sino-Japanese, and Foreign words, from Shibata (1976)

	Gloss				
Stratum	'inn'	'idea'	'acrobatics'	'detour'	'cancellation'
Yamato	宿屋 jadoja	思い付き omoitsɯki	軽業 karɯwaza	回り道 mawarimitʃi	取り消し torikeʃi
Sino-Japanese	旅館 rʲokaɴ	着想 tʃakɯsoː	曲芸 kʲokɯgeː	迂回路 ɯkairo	解約 kaijakɯ
Foreign	ホテル hoterɯ	アイデア aidea	アクロバット akɯrobatto	バイパス baipasɯ	キャンセル kʲaɴserɯ

eign stratum. Almost all Foreign words come from European languages, in particular English, and are generally written using the other moraic writing system in Japanese, *katakana* (Table 1.1), although some very old loans, such as 天ぷら /teɴpɯra/ 'tempura', may be written using *hiragana* or *kanji* instead. Finally, there is a semi-productive set of sound-symbolic words known as the *Mimetic* stratum (Hamano 1998). These are often written using *katakana*, although high-frequency mimetics, such as ばらばら /barabara/ 'scattered' and ぴかぴか /pikapika/ 'sparkle', may be written in *hiragana* instead.

While the lexical strata of Japanese are traditionally defined in terms of their etymology, each stratum also has corresponding phonological, morphological, and stylistic properties distinguishing it from the other strata. Shibata (1976) (cited in Shibatani 1999: 144) gives examples of triplets from the Yamato, Sino-Japanese, and Foreign strata which have roughly the same meaning but vary in stylistic usage (Table 1.2). In general, Sino-Japanese words often have a learned or erudite connotation, and are frequently used in scientific and technical jargon, though many Sino-Japanese words also occur in casual speech, in literary contexts, and in newspaper articles. Foreign terms, on the other hand, have a connotation of stylishness or "coolness", and are often found

Table 1.3: Examples of *rendaku*

X+Yamato	1.	ori	+ kami	→ origami	‘folded paper’
	2.	maki	+ sɯʃi	→ makizɯʃi	‘rolled sushi’
	3.	yama	+ tera	→ yamadera	‘mountain temple’
X+Sino-Japanese	4.	nise	+ kiɴ	→ nisekiɴ	‘counterfeit money’
	5.	oja	+ koːkoː	→ ojakoːkoː	‘filial piety’
X+Foreign	6.	garasɯ	+ keːsɯ	→ garasɯkeːsɯ	‘glass case’
	7.	aisɯ	+ koːhiː	→ aisɯkoːhiː	‘iced coffee’
Lyman’s Law	8.	onna	+ kotoba	→ onnakotoba	‘feminine speech’
	9.	kami	+ kaze	→ kamikaze	‘divine wind’
Nativized	10.	ʧɯːʃakɯ	+ hoɴ	→ ʧɯːʃakɯboɴ	‘annotated book’
	11.	iroha	+ karɯta	→ irohagarɯta	‘*hiragana* playing cards’

in magazines, advertisements, and song lyrics, but less so in newspapers. Both Sino-Japanese and Foreign items tend to be semantically more restricted than Yamato items; for instance, キャンセル /kʲaɴserɯ/ (from English *cancel*) is used specifically to mean cancelling a reservation or an appointment, while 取り消し /torikeʃi/ can be used to mean ‘cancel’ in a more general sense (Shibatani 1999: 144).

1.2.1 *Rendaku* and Lyman’s Law

The Yamato stratum is subject to a morphological process known as *rendaku*, or sequential voicing (Ito & Mester 1986). Basically, *rendaku* causes a word-initial obstruent in the second member of a compound word to be voiced,[4] as shown by the first three examples in Table 1.3. *Rendaku* does not apply if the second member of a compound is

[4] Kuroda (2002) presents an interesting alternative analysis in which the initial obstruent of the second element in these compounds is underlyingly voiced, and is devoiced when it surfaces in word-initial position. For example, the word 紙 /kami/ ‘paper’ in example (1) of Table 1.3 would be represented underlyingly as /gami/; the underlying /g/ surfaces when the word occurs in a compound like 折り紙 [origami], but is devoiced to [k] when it occurs word-initially, resulting in the form [kami].

not Yamato. This is shown in examples (4) and (5) for the Sino-Japanese roots 金 /kiɴ/ ‘money’ and 孝行 /koːkoː/ ‘filial piety’, and in (6) and (7) for the Foreign words ケース /keːsɯ/ ‘case’ and コーヒー /koːhiː/ ‘coffee’. *Rendaku* also does not apply if there is already a voiced obstruent in the second member of the compound. This is shown in (8), where the /b/ in 言葉 /kotoba/ ‘speech’ blocks the /k/ from being voiced, and in (9), where the /z/ in 風 /kaze/ ‘wind’ has a similar effect on the initial /k/. This is known as *Lyman’s Law* (Shibatani 1999: 174). Lyman’s Law also applies within Yamato morphemes, so that, while there are morphemes in the Yamato stratum containing only one voiced obstruent, there are none with two or more voiced obstruents (Ito & Mester 1986).

Finally, there are some words which are historically loanwords, yet still undergo *rendaku*. In (10) the Sino-Japanese root 本 /hoɴ/ ‘book’ becomes [boɴ] in compounds like 注釈本 [ʧɯːʃakɯboɴ] ‘annotated book’,[5] while in (11) the initial /k/ of the loanword カルタ /karɯta/ ‘playing cards’, borrowed from Portuguese *carta*, is voiced in イロハガルタ [irohagarɯta] ‘playing cards with *hiragana* on them’. These words have apparently been nativized to the point where they are seen as members of the Yamato stratum by modern-day speakers (Shibatani 1999: 174).

1.2.2 *NT and *P

There are also various phonotactic constraints governing the lexical strata of Japanese which can interact with morphological processes such as verb conjugation. One of these constraints is a ban against nasals being followed by voiceless stops in the Yamato stratum, which I call *NT, following Ito & Mester (1995). *NT can be seen to operate when

[5] I will explain in the next section, in my discussion of the constraint *P, why it is /h/ instead of /p/ which alternates with /b/ in this form.

Table 1.4: Past tense alternations in Yamato

Gloss	Root	Negative −(a)nai	Gerund −te	Past −ta
‘see’	mi−	minai	mite	mita
‘buy’	kaw−	kawanai	katte	katta
‘die’	ʃin−	ʃinanai	ʃɪɴd̲e	ʃɪɴd̲a
‘read’	jom−	jomanai	joɴd̲e	joɴd̲a

the gerund suffix /−te/ or the past tense suffix /−ta/ is attached to verb bases ending in a nasal, causing the /t/ in both suffixes to be voiced (Table 1.4). This constraint does not hold for the Sino-Japanese stratum, however, and there are numerous examples of post-nasal voiceless stops in Sino-Japanese words, for example, 先輩 /seɴpai/ ‘senior, superior, elder’, 変態 /heɴtai/ ‘transformation; abnormality’, and 文化 /buɴka/ ‘culture, civilization’.

Both Yamato and Sino-Japanese also have a well-known phonotactic constraint on the distribution of the segment [p]. [p] can only appear either as a geminate (as in いっぱい [ippai] ‘a lot, full’), or in a nasal-stop cluster (as in 散歩 [saɴpo] ‘walk, stroll’), but can never appear word-initially or intervocalically. This rather unusual distribution is the result of a sound change which affected word-initial and intervocalic *p in Old Japanese, leniting it to *ɸ. Word-initial *ɸ then changed to /h/, although in many modern-day dialects, /h/ still surfaces as [ɸ] before /ɯ/. Intervocalic *ɸ, however, became *w (again, except before /ɯ/), which was later deleted, except before /a/ (Shibatani 1999: 167). This diachronic relationship between /h/ and /p/ is reflected in several surface alternations between [h] and [p] (or [b]). For example, consider the paradigms for various counter words listed in Table 1.5. 個 /ko/ is a generic counter word, while 冊 /satsɯ/ and 本 /hoɴ/ are specific counters for bound volumes (like books or magazines) and long or cylindrical objects (like pens or umbrellas), respectively. Note that some of the combining forms for numbers, like /iQ−/ ‘one’ and /haQ−/ ‘eight’, when pre-

Table 1.5: Counter words

Number		−ko	−satɯ	−hoN
1	iQ−	ikko	issatsɯ	ippoN
2	ni−	niko	nisatsɯ	nihoN
3	saN−	saNko	saNsatsɯ	saNboN
4	joN−	joNko	joNsatsɯ	joNhoN
5	go−	goko	gosatsɯ	gohoN
6	ɾoQ−	ɾokko	ɾokɯsatsɯ	ɾoppoN
7	nana−	nanako	nanasatsɯ	nanahoN
8	haQ−	hakko	hassatsɯ	happoN
9	kʲɯː−	kʲɯːko	kʲɯːsatsɯ	kʲɯːhoN
10	ʤɯQ−	ʤɯkko	ʤɯssatsɯ	ʤɯppoN

Table 1.6: The /maQ−/ adjectival prefix

Adjective		maQ+Adjective	
kɯɾo	'black'	makkɯɾo	'pitch black'
naka	'center'	mannaka	'dead center'
hiɾɯma	'daytime'	mappiɾɯma	'broad daylight'
hadaka	'naked'	mappadaka	'stark naked'

fixed to a counter, cause gemination of the initial consonant, while other numbers, like /ni−/ 'two' and /go−/ 'five', do not. This is represented in the underlying forms by an abstract segment /Q/ which assimilates in place to a following obstruent. So 'two volumes' is [nisatsɯ], but 'one volume' is /iQsatɯ/→[issatsɯ], with gemination of the initial /s/ of /satɯ/.[6] However, /iQ−/ and /haQ−/, when prefixed to /hoN/, do not result in a geminate [h], but rather a geminate [p]: [ippoN] 'one (pen, umbrella, etc.)'; [happoN] 'eight (pens, umbrellas, etc.)'. A similar pattern appears with the intensifier prefix /maQ−/, which, when attached to adjectives, causes gemination of the initial consonant (Table 1.6). Again, we see an alternation between [h] and geminate [p] in the last two examples.

[6]In the non-loan phonology, underlying /t/ is affricated before /ɯ/, surfacing as [ts].

Table 1.7: Examples of *rendaku* with h → b

ike	+ hana	→ ikebana	‘flower arrangement’
tabi	+ hito	→ tabibito	‘traveler’ (< ‘travel’ + ‘person’)
kake	+ ɸɯtoɴ	→ kakebɯtoɴ	‘top futon’

To account for these surface alternations between [p] and [h], as well as the marked distribution of /p/ in Yamato and Sino-Japanese, Ito & Mester (1995) propose a constraint *P which does not allow an underlying /p/ to surface as [p] unless it occurs as a geminate or in a nasal-obstruent cluster. Under this account, the two cases presented in Tables 1.5 and 1.6 would show an underlying singleton /p/ surfacing as [h] to avoid violating *P. In fact, McCawley (1968) proposes to eliminate the /h/ phoneme from the Yamato and Sino-Japanese strata entirely, deriving all instances of [h] from an underlying /p/. This proposal is not widely accepted, however, because it seems rather counterintuitive to native speakers, and because the change from /p/ to [h] is perhaps an unusual one from a synchronic phonetic standpoint (Shibatani 1999: 167). Of course, diachronically *p > *ɸ > h is an example of lenition, and is not itself an unusual sound change.

*P also interacts with *rendaku*, as the examples in Table 1.7 show. According to Ito & Mester (1995), the initial /b/ present in the second member of 生け花 [ikebana] ‘flower arrangement’, 旅人 [tabibito] ‘traveler’, and 掛布団 [kakebɯtoɴ] ‘top futon’ suggests that there is a /p/ underlyingly in 花 [hana] ‘flower’, 人 [hito] ‘person’, and 布団 [ɸɯtoɴ] ‘futon’, yet again it surfaces as [h] (or [ɸ]) in the latter forms, because a word-initial [p] would violate *P.

Neither *NT nor *P govern the Foreign stratum, though. There are numerous examples of voiceless stops occurring after nasals, as in サインペン /saiɴpeɴ/ ‘sign pen [felt-tip pen]’, インターネット /iɴtaːnetto/ ‘Internet’, and パンク /paɴku/ ‘punk’. And

/p/ is unrestricted in loanwords, occurring not only in geminates or nasal clusters, as it does in Sino-Japanese, but also word-initially (プロセス /pɯrosesɯ/ 'process', ペンパル /peɴparɯ/ 'pen pal') and intervocalically (アパート /apaːto/ 'apartment', タイプ /taipɯ/ 'type').

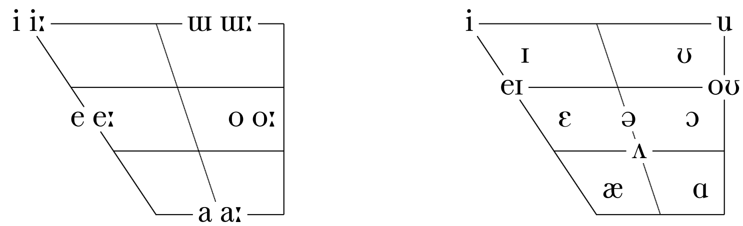
*P can also be violated in the Mimetic stratum. For example, /p/ occurs frequently as the initial consonant in CVCV Mimetic roots (Hamano 1998: 41), as in the words ピカピカ /pikapika/ 'sparkle' or ペラペラ /perapera/ 'fluently'. *NT is obeyed in the Mimetic stratum, however: an obstruent following a nasal within a Mimetic root must be voiced, for example in しょんぼり /ʃoɴbori/ 'being downhearted' and うんざり /ɯɴzari/ 'tedious; boring' (Ito & Mester 1995). This is only true within roots, however, and not across morpheme boundaries, as shown by forms like ピンピン /piɴ+piɴ/→[pimpiɴ] 'lively' or ドスント /dosɯɴ+to/→[dosɯnto] 'with a thump'.

1.3 Japanese loanword adaptation patterns

Non-native segments such as [l], [θ], or [æ] do not occur in Japanese loanwords. While there is a general resistance in the early stages of contact to borrowing non-native segments (Haugen 1950, Thomason & Kaufman 1988), the *katakana* orthography used for loanwords presents an additional challenge to representing these segments in writing. Each *katakana* character represents a single (C)V mora, and unlike alphabetic writing systems, there is no systematic relationship between the form of a character and the segment(s) that it represents, making it nearly impossible to indicate foreign segments without introducing completely new characters for them, or at least applying diacritics to previously-existing characters.[7] This poses a problem for adapting words from English,

[7] This latter possibility was employed in the creation of the *katakana* character ⟨ヴ⟩, which is the character ⟨ウ⟩ /ɯ/ with a *dakuten* (voicing mark) added to it. ⟨ヴ⟩ is used in loanwords to indicate a

Table 1.8: Vowel systems of Japanese and American English



(a) Japanese (b) English

which has a larger segment inventory than Japanese (Tables 1.8 and 1.9). The convention for adapting English vowels that has arisen over time is that the tense-lax distinction in English is reinterpreted as a length distinction in Japanese. English lax vowels are adapted into Japanese as short vowels, while English tense vowels are adapted as Japanese long vowels, for example in *bit*→/bi̲tto/, *beat*→/bi̲ː̲to/ (Katayama 1998; see also section 3.3.1 in Chapter 3). The English diphthongs /ɑɪ/, /ɑʊ/, and /ɔɪ/ are adapted into Japanese as sequences of two short vowels, /ai/, /aɯ/, and /oi/, respectively.

The Foreign stratum also has certain phonotactic constraints governing it which affect how loanwords are adapted. Generally speaking, these constraints are a superset of the constraints governing the other three strata. For example, a generalization governing both native words and loanwords is CODACOND (Ito & Mester 1995) or *CODA (Katayama 1998), which restricts syllable codas to being either a moraic nasal or the first segment of a geminate. Thus word-final consonants from English source words must be repaired in some way to avoid violating CODACOND. The usual strategy used in recent loans is to preserve the word-final consonant via epenthesis.[8] Katayama (1998) explains the occurrence of epenthesis in loanwords by ranking the constraints *CODA and MAX-IO (no deletion) above DEP-IO (no epenthesis), as in Table 1.10. However, there

source word /v/, as in ラヴ /ɾabɯ/ ‘love’. Yet ⟨ヴ⟩ is normally pronounced /bɯ/, and words with ⟨ヴ⟩ are often spelled with the corresponding /b/ *katakana* instead (for example, ラブ for *love*).

[8]The epenthetic vowel usually used is /ɯ/, but /i/ is used for word-final /ʧ/ or /ʤ/ (but not /ʃ/), and /o/ for /t/ or /d/.

Table 1.9: Consonant inventory of Japanese. Segments in parentheses occur only as non-contrastive allophones (except /ɸ/ and /ts/, which are contrastive in the Foreign stratum only), while segments with an asterisk occur only in the Foreign stratum. The segments /tʃ/, /dʒ/, and /ʃ/ actually have an alveolopalatal place of articulation (Akamatsu 1997), but in keeping with other studies of Japanese phonology I transcribe them as palatoalveolars. The palatal segments /tʃ/, /dʒ/, /ʃ/, and /ç/ contrast with their non-palatal counterparts /t/, /d/, /s/, and /h/ only before back vowels in Yamato, Sino-Japanese, and Mimetic words. These palatals are often analyzed as underlying /tʲ/, /dʲ/, /sʲ/, and /hʲ/ (Ito & Mester 2003).

p	pʲ	t	*tʲ			k	kʲ	
b	bʲ	d	*dʲ			g	gʲ	
		(ts)		tʃ				
		(dz)		dʒ				
(ɸ)	*ɸʲ	s		ʃ	ç			h
		z						
m	mʲ	n	nʲ	(ɲ)		(ŋ)		(ɴ)
w		ɾ	ɾʲ		j			

Table 1.10: Adaptation of *bike*→[baikɯ], based on Katayama (1998)

	baɪk	*CODA	MAX-IO	DEP-IO
a.	bai		*!	
b.	baik	*!		
c. ☞	bai.kɯ			*

Table 1.11: Possible voiceless coronal-vowel and palatal-vowel sequences in Japanese. Sequences in parentheses are found only in loanwords, while those marked with a star are rarely attested at all.

(ti)	*si	*tɯ	sɯ	ʧi	ʃi	ʧɯ	ʃɯ
te	se	to	so	(ʧe)	(ʃe)	ʧo	ʃo
		ta	sa			ʧa	ʃa

Table 1.12: Palatalization in Japanese verb alternations

Gloss	Root	Negative −(a)nai	Polite −(i)masɯ	Past −(i)ta
'see'	mi−	minai	mimasɯ	mita
'wait'	mat−	matanai	maʧimasɯ	matta
'lend'	kas−	kasanai	kaʃimasɯ	kaʃita

are some cases of older loans where deletion of the illegal coda consonant occurs instead. For example, the mid-19th century borrowing *lemonade*→[ɾamɯne] shows deletion of the word-final /d/ (compare with the more recent reborrowing of *lemonade* as [ɾemoneːdo], where the source word-final /d/ occurs in the loanword with an epenthetic [o]).

There are also many examples of words in the Foreign stratum which have phonotactic sequences which do not occur at all in the other three strata. For example, coronal obstruents can occur before the vowel /i/ in some loanwords, for example パーティー /paːtiː/ 'party' and ディレクター /diɾekɯtaː/ 'director'. However, coronals never surface before /i/ in non-Foreign words (Table 1.11). In cases where a coronal-/i/ sequence is generated underlyingly, such as in verb conjugation patterns (Table 1.12), the coronal is palatalized. There are also some older loans, such as ジレンマ /ʤiɾeɴma/ 'dilemma', which show a palatal-/i/ sequence corresponding to a coronal-/i/ sequence in the source word. Ito & Mester (1995) have suggested that palatalization in both cases results from the interaction of faithfulness constraints with a markedness constraint *TI

Table 1.13: Adaptation of *dilemma*→[ʤireɴma]

	dɪlɛmə	*TI	FAITH
a.	direɴma	*!	
b. ☞	ʤireɴma		*

(coronal stops may not occur before /i/), as shown in Table 1.13.

The analyses above of epenthesis and palatalization in loanwords are typical of the prior research on loanword adaptation. In long-term contact situations between two languages L1 and L2, L1 speakers will develop over time a set of specific conventions for borrowing words from L2. These conventions can be stated in the form of a set of rules (or ranked constraints, under an OT approach) transforming an L2 representation of the source word into an L1 phonological representation, resembling the generative conception of the phonology of a language as a set of rules (or constraints) transforming an abstract underlying representation into a surface representation that can be phonetically interpreted. This resemblance is perhaps what has led many researchers to make the simplifying assumption that nativizations are performed by a single speaker (presumably a bilingual, or an L1 speaker with access to L2 speakers), and that these nativizations reflect the effects of the shared phonological grammar common to all L1 speakers. Nativizations seen in established loanwords are thus explained as resulting from the process by which a single borrower, in the perception and/or production of a new loanword, maps the non-native segments and phonotactic sequences in the source word to phonemes and licit sequences in her native language (Silverman 1992, Yip 1993).

However, this general framework for analyzing adaptation patterns, while successful in describing the established system of conventions used by modern-day Japanese speakers in borrowing words from English,[9] is hard to reconcile with the existence of

[9]Ito & Mester (1995) and Katayama (1998) offer recent OT accounts of Japanese adaptation patterns,

both synchronic and diachronic variation in adaptation patterns. Early loanwords that are borrowed at the beginning stages of contact between L1 and L2 typically show a variety of repair strategies for non-native patterns. Over time, as the level of bilingualism in a speech community increases, this initial variation coalesces to a single agreed-upon convention for adapting non-L1 patterns from L2 (Haugen 1950). As was mentioned above, some older loanwords like /ɾamɯne/ (borrowed from English *lemonade*) show deletion of illegal word-final consonants, yet in more recent loanwords these consonants are instead preserved via epenthesis.[10] Likewise, in older loans, coronals before /i/ are generally palatalized, yet in recent loans, coronals are instead preserved before /i/, introducing a new contrast between coronals and palatals in this environment. There are many other examples like this of changes over time in adaptation strategies in Japanese loanwords. The question is, why did these changes take place, and how were they implemented in terms of the actions of individual speakers over time? These questions, much like the analogous questions concerning sound change in general (Weinreich et al. 1968, Lass 1980), are difficult to address under an approach which looks only at the behavior of a single idealized speaker at a particular moment in time.

The assumption that the adaptations seen in loanwords can be explained in terms of the phonological grammar of a single speaker is also questionable when we examine the evidence for synchronic variation. Even in cases of long-term contact, loanwords which are only used by a few L1 speakers typically show less nativization than those which are used by a larger number of speakers (Poplack, Sankoff & Miller 1988). This suggests that it is not only the initial borrowers of a loanword, but also the L1 speakers involved in transmitting the loanword to the broader speech community, who have a role in performing loanword adaptations. The establishment of loanword adaptation

while Ohso (1971) and Lovins (1975) provide earlier accounts using the framework of Natural Phonology (Stampe 1973).

[10] Smith (2006) analyzes these cases as being based on different types of input (spoken or written loans), as I will discuss in section 3.1 of Chapter 3.

patterns is thus the result of the cumulative action of many speakers operating over generations of contact with L2, and requires a theoretical approach which can model both the phonological competences of individual speakers, and the interactions among these speakers within the broader speech community.

1.4 Outline for the remainder of the dissertation

The rest of this dissertation will be structured as follows. I begin in Chapter 2 by looking at the process of loanword borrowing, analyzing it as two separate processes: the initial *adaptation* of a new loanword by one or more speakers of a speech community, and the subsequent *transmission* of the borrowed word to other speakers. I argue that nativization of a loanword can be done not only by the initial borrowers of the word (as previous approaches have assumed), but also by the speakers involved in transmitting the word. This allows for variation in attested adaptation patterns which cannot easily be accounted for by approaches that look at the process of adaptation only. I then present historical data showing how adaptation patterns in Japanese loanwords have changed over time, focusing on three processes: the palatalization of velars before a source word /æ/, gemination of word-final voiced obstruents, and palatalization of coronals before front vowels. It turns out that these processes show varying degrees of regularity in loanword adaptation, depending on both the time of borrowing, and (in some cases) the type frequency of the phonological neighborhood that the non-native pattern appears in. Generally speaking, the longer a word is attested for, the more likely it is to be nativized, while the preservation of non-native elements is more likely in phonological neighborhoods that are of higher type frequency among the set of loanwords in the language. Two important dates, 1870 and 1890, corresponding to the opening of Japan to the Western world, and the establishment of mandatory English education, respectively, will be seen to coincide with changes in adaptation patterns, with most synchronic variation in adaptations ceasing after about 1890.

Chapter 3 looks more closely at the process of loanword adaptation by a single speaker. After examining the debates over whether the input to borrowing is a phonological representation or not, and whether adaptations take place during perception or production, I then turn to the issue of how loanwords are represented in a speaker's lexicon. The traditional answer to this is that loanwords are marked as such by being stored in a special lexical stratum separate from the native items of the lexicon. I examine two recent OT proposals for representing lexical stratification, the *Core-Periphery* model (Ito & Mester 1995) and the *Cophonology* model (Inkelas & Zoll 2007), and show that neither model can provide an explanation for the influence of type frequency and phonological neighborhood on the likelihood of nativization.

In Chapter 4 I present an alternative conception of the lexicon, based on connectionist research on word reading. I begin by introducing connectionist models of language processing, arguing that they can provide an explanation for frequency and similarity effects, which arise naturally from the learning and processing mechanisms of such models. I review Seidenberg & McClelland's (1989) connectionist framework for modeling single-word reading, and consider how to apply it to the problem of explaining loanword adaptations. I then design and analyze two connectionist models in this chapter. The first model is a simple feedforward network which is trained to identify the lexical stratum of words presented to it. I use the behavior of this network to show that lexical strata need not be represented explicitly in the minds of Japanese speakers, but rather can be thought of as emergent structures in the lexicon, composed of smaller-scale phonological neighborhoods. I argue that it is membership in these phonological neighborhoods, and not stratal membership, which drives nativization at the level of the individual speaker, allowing for more fine-grained adaptation strategies sensitive to the type frequency of the phonological neighborhood a non-native pattern occurs in. I then develop another connectionist network which is trained to replicate an input pattern of phonological features representing a word, repairing invalid feature values as necessary. I use this network to examine the voiced geminate pattern from Chapter 2, showing that the relative acceptability of different voiced geminates arises to some extent from the frequency of the corresponding voiceless geminates in the native lexicon.

In Chapter 5 I shift focus from the individual borrower to the speech community as a whole, as I look at how to characterize the effect that transmission has on the nativization of loanwords. Although the discussion in this chapter is reminiscent of sociolinguistic work using network methodologies, such as that of Milroy (1987), it turns out that social network theory, with its focus on individual actors and the roles they play within small-scale networks on the order of a dozen or so nodes, cannot be directly applied to the historical data presented in Chapter 2, for which we have little if any information about the social relationships among the Japanese speakers who produced the attested loanwords. Instead I develop a broader framework, based on the recent physics literature concerning the large-scale structural properties of social networks (Watts & Strogatz 1998, Albert & Barabási 1999), for modelling the effects of transmission within an idealized social network. I show that the main effect of transmission is to amplify the effect of nativization at the individual speaker level, so that even if speakers individually have only a relatively small tendency to nativize a particular loanword, this can still result in the speech community as a whole adopting the nativized form of the word. I then apply this model to the data from Chapter 2, showing that, before about 1890, the nativization of coronals before /i/ by individual speakers seems to have taken place at a qualitatively greater rate than the nativization of other non-native phonotactic patterns. I suggest that this is because both phonological-level processes (i.e. a markedness constraint against /ti/, /di/, etc.) and phonetic-level processes (misperception and/or misproduction of [ti], [di], etc.) were involved in the nativization of coronal-/i/ sequences, whereas nativization of other sequences involved phonetic-level processes only. After 1890, due perhaps to increased English education or to an increase in the number of new loanwords being borrowed into the language, the grammatical constraint banning coronal-/i/ sequences seems to have been no longer active for loanwords for most Japanese speakers, and nativization by individual speakers at this time dropped to roughly the same rate as nativization of other phonotactic patterns.

Finally, in Chapter 6 I summarize the main results from the previous chapters, propose extensions to the models developed in Chapters 4 and 5, and provide directions for future research.

1.5 A note on terminology

The terminology used in loanword studies is often quite confusing, with different authors using words like "borrowing" and "adaptation" in very different senses. For the sake of clarity, I am going to be defining these terms in the remainder of this dissertation as follows:

Borrowing

The process by which a new loanword becomes established in a speech community, involving both the *adaptation* of the loanword by one or more L1 speakers, and the subsequent *transmission* of the loanword from these initial borrowers to the other members of the speech community.

Adaptation

The process of mapping an L2 input to an L1 loanword. The L2 representation may be phonetic, orthographic, or phonological in nature. Note that many authors use the term "borrowing" to refer indiscriminately to both adaptation and borrowing (as I have defined them here).

Transmission

The spread of a loanword throughout a speech community. As I will argue in Chapter 2, nativizations (defined below) can take place during the transmission stage of borrowing, in addition to the adaptation stage.

Nativization

A change made in the phonological form of a loanword to make it conform more closely with L1 phonology and phonotactics, at the expense of more faithfully representing the L2 source word.

CHAPTER 2

ADAPTATION AND TRANSMISSION OF LOANWORDS

2.1 The borrowing process

What exactly do we mean when we say that a language L1 has borrowed a word from another language L2? Ultimately, since languages are not themselves atomic entities, but rather are convenient labels for the aggregate linguistic behaviors and mental states of the members of particular speech communities, this statement must reduce to a description of a change in the behaviors and mental states of the individual speakers of L1.[1] Thus when we talk of a word being borrowed from L2 into L1, what we are referring to is the process by which a large enough number of L1 speakers have, over time, added the word to their mental lexicons so that the word is generally recognized as being an element of the L1 lexicon, to the point where it can be used in monolingual L1 contexts. This definition of borrowing makes it clear that it is really a change in the L1 speech community overall, and not just a change in a single speaker's L1 lexicon, which is important for borrowing, although the change in the speech community is composed of many individual changes in speakers' lexicons. If only one L1 speaker has added a new L2 word to her L1 lexicon, then we cannot really speak of borrowing in this case, since other L1 speakers will not necessarily understand her if she uses the word in an L1 context. At most, we can call this a potential borrowing, if usage of the new word eventually spreads to other L1 speakers.

[1] My definition of language here as an aggregate construct is somewhat analogous to the *populationist* definition of species in biology (Mayr 1982; see also Ghiselin 1974, 1997), in which a species is defined as a population of individual organisms that can potentially mate with each other and produce viable offspring. This is in contrast with the *essentialist* definition of species as idealized types with a set of common properties, which resembles the more common conception of languages as abstract entities defined by their grammatical properties. Essentialist notions of species make it difficult to talk about how one species evolves into another over time; likewise, essentialist notions of language make it difficult to talk about language change.

From this viewpoint of looking at loanword borrowing as a change in the speech-community as a whole, we can identify two important stages in the establishment of a new loanword: the initial *adaptation* of the word by one or more L1 speakers who have some sort of exposure to L2 (through bilingualism, or access to L2 speakers or texts), and the subsequent establishment of the word in the overall L1 speech community via its *transmission* from the initial borrowers to other L1 speakers (who may or may not also have contact with L2). This is similar to the distinction between language *change* in itself, and *diffusion* of such changes, made by Hale (2007). In fact, Hale mentions the replacement of Middle English *lutter* with the French loanword *pure* as an example of diffusion (p. 39).

The adaptation and transmission stages need not be mutually exclusive. It is possible for later adaptations of the same L2 word to occur while transmission of earlier adaptations is already taking place among other speakers. This can sometimes cause interference effects, if two different adaptations of the same L2 word are being transmitted through the speech community, as I will show later in the discussion of palatalization of velar stops before /æ/ in section 2.3.3. While *nativizations* (changes in the phonological form of a loanword to avoid violating L1 markedness constraints, at the expense of faithfulness to the original L2 source word) are normally thought to be made during the adaptation stage, they can arise during the transmission stage as well, as I will argue later on.

The distinction I am making here between the adaptation and transmission stages of loanword borrowing[2] is analogous to the distinction Poplack et al. (1988) make between *nonce* borrowings (borrowings that only occur once in their corpus) and *widespread* bor-

[2]As I mentioned at the end of Chapter 1, I am using the term *borrowing* only to refer to the overall process by which a loanword is established in a speech community, and which is composed of both the adaptation and transmission stages.

rowings (borrowings that were used by more than 10 speakers),[3] as well as Shinohara's (1996) distinction between *phonemicizations* and *loanwords*:

> Phonemicizations are, in principle, different from loanwords in that one can learn loanwords without being exposed to the source language. Loanwords are already in the Japanese lexicon and they can be transmitted from one Japanese speaker to another. A phonemicization is something created on the spot. (p. 90)

However, other studies of loanword phonology are not as careful as Poplack et al. (1988) or Shinohara (1996) in distinguishing between the adaptation and transmission stages of loanword borrowing. Instead, they focus on the adaptation stage only, and use the nativizations seen in loanwords to reason about the phonological competence of the L1 speaker(s) who originally borrowed the word:

> Loanword adaptation is constraints and repairs in "real time". In adapting a loan the speaker tries to remain faithful to the source word while still making the loan conform to the native language (L1) segmental inventory, phonotactic constraints, and prosodic structures. Because inputs of considerable diversity and complexity can be devised, loanword phonology takes on the status of something akin to an "experiment of nature" in allowing us to probe phonological competence. (Kenstowicz & Suchato 2006: 921)

Kenstowicz & Suchato (2006), and many other researchers, do of course acknowledge that there are other factors as well that are involved in determining the ultimate phonological form of an established loanword. For example, following the previous quote,

[3]Poplack et al. (1988) also distinguish between *idiosyncratic* (used by only one speaker) and *recurrent* (used more than 10 times, but not necessarily by 10 different speakers) borrowings. As they note, these latter two types subsume the nonce and widespread categories of borrowings, respectively.

they continue by saying, “Needless to say, various methodological issues have arisen in this enterprise such as *the distinction between on-line versus integrated loans*, the role of orthography, as well as the often-variable nature of the data...” (p. 921, my emphasis) Yet the theories of loanword borrowing that they have developed consider only the role of the phonological competence of a single speaker performing an adaptation of a new loanword. This is of course not a problem if we are only trying to explain individual speakers’ on-line adaptations, as in experimental studies like Dupoux et al. (1999) and Vendelin & Peperkamp (2006), in which speakers are presented with a novel L2 source word and then asked to produce an L1 adaptation of that word. However, many loanword studies are trying to account for nativizations occurring in established loanwords as well. For example, Kenstowicz & Suchato (2006) use an 800-word corpus of English loanwords borrowed into Thai, 90% of which were collected from an English-Thai dictionary (and can thus be considered to be established loanwords), and the rest of which were collected from Thai students in the United States (which are probably a mixture of nonce loanwords, established loanwords, and loanwords which are in the process of becoming established). Likewise, Ito & Mester (1999) use data both from established loanwords and native speaker judgments of possible adaptations to support their arguments for a stratified lexicon. Silverman (1992) uses corpus data, words elicited from native Cantonese speakers, and experimental data from a forced judgment task where native speakers had to choose which of several possible loanword forms would be the best adaptation of a given English word, while Yip (1993) rules out entirely what she calls “unassimilated” loans,[4] so that her data consists solely of established loanwords.

The data used in typical loanword studies thus consists largely of established loanwords, which are the end product of both the adaptations of the L2 source by the initial L1 borrowers, and the subsequent transmission of these adaptations to other L1 speak-

[4]These seem to be loans which contain syllables that do not occur in native Cantonese words (these syllables being either accidental gaps, or ones ruled out by Cantonese syllable structure constraints).

ers. However, this data is used to reason about the adaptation process only. This makes sense only if we assume that any nativization attested in a given loanword must originally be the result of an adaptation from an L2 source performed by a single speaker, and that the loanword was then faithfully transmitted in this form to other speakers, who did not make any further changes themselves to the loanword. In the case of borrowings into present-day Japanese, this is a reasonable assumption to make. Japanese has been in contact with English for about 150 years, and over that time Japanese speakers have developed a set of well-known conventions for adapting words from English that are applied consistently to recent loans. An example is the adaptation of English /l/ and /ɹ/, neither of which exist in the phonemic inventory of Japanese. /l/, in both onset and coda position, is always adapted into Japanese as /ɾ/, for example in ライブ /ɾaibɯ/ 'live' and アイドル /aidoɾɯ/ 'idol'.[5] /ɹ/, however, is adapted as /ɾ/ only in onset position, for example, in ライス /ɾaisɯ/ 'rice'. Coda /ɹ/ is adapted as either /a/, when it occurs word-finally (as in コア /koa/ 'core'), or is deleted while triggering lengthening of the preceding nucleus when it occurs word-medially (as in コース /koːsɯ/ 'course'). These conventions for adapting /l/ and /ɹ/ are well-established among Japanese speakers, and are applied consistently to more recent loans from English such as インターネット /iɴtaːnetto/ 'Internet'. Thus it is conceivable that, for modern-day speakers, these conventions are a part of their phonological grammar, and can be applied by any speaker on a never-before-seen English word to produce the correct adaptation if it were to be borrowed into Japanese. But the same is not necessarily true of historical Japanese speakers at the time that conventions like those governing the adaptation of /l/ and /ɹ/ were being established. In fact, very early loans that date from before adaptation conventions were established will often show quite a bit of variation in attested adaptation patterns. For example, words like チェア /ʧea/ 'chair' and ドア /doa/ 'door', both of

[5] /aidoɾɯ/ also has an epenthetic /ɯ/ following the adaptation of /l/ as /ɾ/, since Japanese does not allow non-nasal coda consonants in word-final position.

which date from the 1860's according to Arakawa (1977), are attested as チェル /ʧeɾ̠ɯ/ and ドル /doɾ̠ɯ/, respectively, before the /ʧea/ and /doa/ forms became the established ones. These two early forms show word-final /ɹ/ adapted as /ɾ/, rather than /a/, as would be expected for more recent loans. Modern-day adaptation patterns are the end result of the collective decisions of many speakers historically over several generations, and so it is not appropriate, when talking about how these adaptation patterns became established, to model them solely in terms of the behavior of an individual speaker. It is also necessary to consider how the interactions between speakers over time led to the establishment of a single adaptation convention used by the future members of the speech community for later borrowings.

Given the discussion above, there is reason to believe that, beyond any initial nativizations made during adaptation by initial borrowers, further nativizations can be made by the speakers involved in transmitting a loanword as well, as I will now discuss in the next section.

2.2 Nativization during transmission

I will first make a theoretical argument, following Hale & Reiss (2000, 2001) and Blevins (2004), that markedness effects seen in loanword adaptations may be due to the cumulative production and perception biases of speakers as a loanword is transmitted among them.[6] Consider the case discussed in section 1.3 of Chapter 1 of the adaptation of coronal obstruents before the high front vowels /i/ and /ɪ/ (both of which are adapted into Japanese as /i/). In older loans, at least, source-word coronals are palatalized in

[6]Blevins (2004) considers how to explain language change as a result of the accumulation from generation to generation of perceptual biases during language acquisition. Hale & Reiss (2000, 2001) argue for a similar mechanism for language change, while focusing more on the role of Universal Grammar.

Table 2.1: Adaptation of *team*→[tʃiːmɯ], *[tiːmɯ]

tim	*TI	FAITH
a. tiːmɯ	*!	
b. ☞ tʃiːmɯ		*

this environment, for example in the word チーム /tʃiːmɯ/ 'team'. Since there is a similar process of palatalization of coronals before /i/ in the native phonology, Ito & Mester (1995, 1999) have suggested that both processes are triggered by the same set of markedness constraints *TI and *SI forbidding coronal stops and fricatives, respectively, before /i/.[7] In other words, when a Japanese speaker tries to borrow a word like *team* from English, since *TI ≫ FAITH, they will adapt the illegal /ti/ sequence from English as [tʃi] instead of [ti] in the resulting loanword (Table 2.1).

However, let us consider what would happen if the speaker were to instead adapt *team* as [tiːmɯ], and then this form were to spread among other Japanese speakers who are biased towards perceiving [ti] as [tʃi], but do not necessarily have a categorical constraint like *TI forbidding all [ti] sequences. Assume there are N speakers who are arranged in a line like in the game of "Telephone", where each person hears the word spoken by the person on their left, and then repeats it to the person on their right (Figure 2.1). Suppose that each speaker, with probability p, will correctly reproduce a [ti] sequence in the word they hear, and with probability $1-p$ will either misperceive or misproduce the [ti] sequence as [tʃi]. Then, the likelihood that the Nth speaker will produce [tiːmɯ] will be $p \times p \times p \times \ldots = p^N$, and thus the likelihood of [tʃiːmɯ] will be $1-p^N$. If $p<1$, then p^N will tend to decrease with increasing N, and likewise $1-p^N$ will tend to increase. In other words, as the number of speakers increases, the likelihood

[7]Ito & Mester need to propose two separate constraints in order to account for the fact that in more recent loans, only coronal fricatives are palatalized in this environment, as in the loanword シティバンク /ʃitibaɴkɯ/ 'Citibank'. Native phonological processes do not distinguish between stops and fricatives in this environment, however; both are palatalized when they occur before /i/.

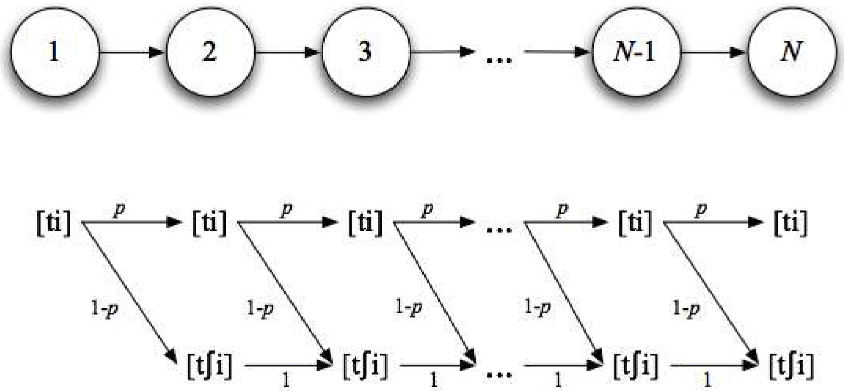


Figure 2.1: The game of Telephone played with loanwords. Each speaker (represented by the circles) hears a word produced by the speaker on their left, and then tries to reproduce the word for the speaker on the right. The first speaker always produces [ti]. Each following speaker, if the last speaker produced [ti], will also produce [ti] themselves with probability p, or will instead produce [ʧi] with probability $1-p$. However, if the last speaker produced [ʧi], then the next speaker will always produce [ʧi] as well. The probability that the final speaker will produce [ti] is p^N, and the probability that he will produce [ʧi] is $1-p^N$.

that the final speaker in the chain will produce [ʧi] instead of [ti] also increases. For example, with $p = 0.8$ and $N = 10$, the probability that the Nth speaker will produce [tiːmɯ] is only $0.8^{10} \approx 0.11$, while with $N = 20$, this probability drops to $0.8^{20} \approx 0.01$. Of course real-world speech communities have a far more complex structure than a simple linear chain of speakers, as I will discuss in chapter 5. What this simple model does show, though, is that the existence of palatalization in established loans like [ʧiːmɯ] can be the result not only of the effect of categorical constraints during adaptation, but also of the cumulative effect of misperceptions and misproductions during transmission as well.

Of course it is difficult to know to what degree the nativizations seen in Japanese loans like [ʧiːmɯ] is due to constraints during adaptation or to cumulative misperceptions/misproductions during transmission without detailed empirical data showing precisely who introduced a given loan when, and how usage of that loan then spread from speaker to speaker. One of the few studies to approximate this type of data is Poplack et al.'s (1988) study of English borrowings among French speakers in the Ottawa-Hull area. They show that the degree to which a given loan is nativized (measured by an index of integration, which ranges from 0 for a word rendered completely using English phonology, to 1 for completely in French phonology) depends on the number of speakers using the loan in their corpus:

> ...phonological integration proceeds as a function of the social integration of the loanword...nonce borrowings are about as likely to be rendered in English as in French, with only a slight bias in favor of the former. Compare this with code-switches into English, which receive English phonology three-quarters of the time. As we move to the most widespread words, the index of integration rises steadily, so that the likelihood of words used by over 20 speakers receiving French phonology is very high, and English phonology very low. (pp. 72–3)

Poplack et al. show that there is a correlation between the number of speakers using a loanword and that loanword's index of integration. There are at least a couple of possible mechanisms taking place during the process of loanword transmission which would produce this correlation. One possibility, given the results above from my simple model of loanword transmission with errors, is that as a loanword is transmitted among more and more speakers, in the process it also becomes more and more nativized as compared to nonce borrowings produced by a single speaker. In other words, the speakers transmitting the loanword are not simply mimicking the loan in exactly the form that they learned it, but are occasionally performing some nativizations of their own as well. Another possibility is that the French speakers in the study were more likely to acquire and use a new loanword they have heard in the speech of their neighbors in their social network if the word already had a relatively high index of integration. Then it would be easier for highly-nativized words to spread among more speakers than less-nativized words. Both of these processes could work in tandem to produce the correlation between the number of loanword users and the index of integration. Then, to the degree to which the first mechanism (nativization by transmitting speakers) is active during the transmission of a loanword, the correlation between the degree of nativization and the number of speakers using a loan would be indirect evidence for the effects of misperception and misproduction during loanword transmission on the likelihood that the nativized form of a loanword will become the established form in a speech community.

More direct evidence for transmission effects on loanwords in Japanese can be seen in a particular set of loans from English, namely those derived from source words containing a velar stop followed by /æ/. In some cases, velars in this environment will be palatalized in the resulting loanword, for example in キャンデー /kʲaɴdeː/ 'candy' and キャット /kʲatto/ 'cat'. This palatalization seems to be a reflection of the relatively fronted articulation of velars before front vowels in English (Keating & Lahiri 1993). Yet in other loans, such as カテゴリー /kategoriː/ 'category' or カメラ /kamera/ 'camera', the velar is not palatalized. In previous work (Crawford 2004) I suggest that this variation is essentially random, due to /æ/ being less front than /i/ and thereby tending to cause less fronting on the preceding velar, making it possible that in some cases a Japanese listener will perceive the velar as a plain stop instead. However, it turns out that, for loans which are first attested after about 1890, the variation is largely dependent on a single factor: whether the English source word has a transparent cognate in French or German.[8] If there are no transparent cognates, then the loan will almost always have a palatalized velar stop; however, if cognates do exist, as in *category* and *camera*, which

[8]For earlier loans, the situation is somewhat more complicated, since there is a great deal of variation in adaptation patterns of loans attested between roughly 1850–1890, with [ka], [kʲa] and [ke] all occurring as possible adaptations of English /kæ/, even in loans that do not have transparent cognates in other languages.

are cognate to the French *catégorie* and the German *Kategorie* and *Kamera*, then the loan is more likely to have a plain velar stop instead.[9]

I will be discussing these adaptation patterns in more detail in section 2.3.3, but the important point here is that there is no satisfactory way to explain why a velar in a given loanword is palatalized or not without making reference to the process of loanword transmission. Since predicting the correct outcome depends on knowing whether cognates exist in French or German, any account which tries to explain velar palatalization patterns solely as the result of a single speaker performing an adaptation will have to make the highly unlikely assumption that all of the Japanese speakers who originally borrowed these words were familiar enough with all three languages (English, French, and German) that they knew that words like *category* have transparent cognates and should be borrowed with a plain velar, whereas words like *cat* only exist in English and should be borrowed with a palatalized velar. A transmission-based account, on the other hand, does not need to assume that the borrowers of these words must have been in contact with all three languages. Instead, we can make the more reasonable assumption that when a word like *category* was originally introduced into the Japanese speech community, speakers who were familiar with English would have adapted the English word, producing [k^jategoriː] with a palatalized /k/, while other speakers who were more familiar with French or German would have adapted the French or German word, producing [kategoriː] with a plain /k/. As both [kategoriː] and [k^jategoriː] then spread through the Japanese speech community, there were some speakers who were exposed to both forms and had to choose between using one or the other in their own speech. If we assume

[9] Of course, *cat* is cognate to French *chat* and German *Katze*, yet the /k/ is still palatalized in [k^jatto], so it seems that *chat* and *Katze* are not similar enough to *cat* to block velar palatalization in this case. What I mean here by a transparent cognate is a cognate similar enough that an adaptation of the French or German word would be identical segment-for-segment to the adaptation of the English word, except for the palatalization on the velar stop. Thus the word *chat*, if borrowed into Japanese, would come out as something like [ʃaː], which is too dissimilar from [k^jatto] to have any effect on the palatalization of the velar.

that these speakers tended to pick the version with the plain stop over the palatalized stop (perhaps because palatalization is more difficult from an articulatory standpoint), then [kategoriː] will gradually become more common than [kʲategoriː] during the process of loanword transmission, and will eventually become the established form of the loanword. However, with a word like *cat*, there are no alternate adaptations from French or German with plain /k/ competing with the adaptation from English with palatalized /k/ as it spreads through the speech community, and so [kʲatto] will become the established form by default. In a sense, then, the knowledge that a particular word has a transparent cognate does exist in the speech community as a whole, but does not need to be localized in any particular speaker in order to affect adaptation patterns. Instead, this knowledge manifests in the competition between palatalized and unpalatalized variants during the process of loanword transmission, resulting in an adaptation pattern that cannot reasonably be attributed to the action of a single speaker.

To summarize, while nativizations are commonly assumed to be the result of loanword adaptation only, it turns out that further nativizations can be made by speakers involved in transmitting a loanword as well. Thus we must be careful about using nativization patterns to probe into the phonological competence of individual speakers. Specifically, markedness effects seen in nativizations may not always be due to markedness constraints operating in a single speaker's phonological competence, but can also result from cumulative perception and production biases during the process of loanword transmission. Empirical studies of loanword borrowing, like Poplack et al. (1988), show that the level of nativization in a loanword depends on the number of speakers using the word, which suggests that nativizations are being performed not only by the initial borrowers of a word, but also by those speakers involved in transmitting the word as well. In other words, nativization is not a single discrete event involving the initial borrowers only, but is instead a gradual process involving the actions of many different speakers in addition to the initial borrowers. Finally, the facts concerning palatalization of velars before /æ/ in Japanese loanwords are most satisfactorily explained as arising from the competition of cognate loans during transmission, and cannot be explained at all as arising solely from the initial adaptation without attributing unrealistic knowledge of the existence or not of cognates to the original borrowers of these words.

2.3 Changes in adaptation patterns in Japanese loanwords

In this section I will present data on how foreign words containing various phonotactic patterns unattested in the Yamato or Sino-Japanese strata of the Japanese lexicon have been adapted into Japanese, and how these adaptation patterns have changed over time. I will examine the following three source patterns: velar stops before /æ/, which, as we saw above, may or may not be palatalized in the resulting loanword; word-final voiced stops following a lax vowel, which can be geminated, despite voiced geminates not occurring in native words; and the distribution of coronal and palatal obstruents before the front vowels /i/ and /e/, of which only palatal-/i/ and coronal-/e/ sequences can occur in the native phonology.

2.3.1 Loanword-specific faithfulness constraints

For each set of adaptation patterns that follows, I will present data collected from Arakawa (1977) showing how it was attested in various loanwords over time, then I will discuss a possible OT analysis of the attested adaptations. While it is generally useful to express the various factors at play in determining an adaptation pattern using a set of ranked OT constraints, in order to adequately characterize the historically attested adaptation patterns I will need to reject two assumptions that are commonly made in the OT literature on loanword borrowing. The first assumption is that there are no loanword-specific constraints operative in the grammar of the borrowing language (Yip 1993). Specifically, the IO family of faithfulness constraints (Kager 1999) cannot be used to model the interactions between faithfulness and markedness in adaptations, because the repair strategies seen in loanwords may sometimes conflict with those used in the native phonology (Kang 2003, Yip 2006). Smith (2006) points out that while epenthesis is normally used to repair illegal coda consonants and consonant clusters in Japanese loanwords (for example, in *best*→[besɯ̲to̲]), deletion is used instead for repairs in verb conjugation patterns (/jom+ɾɯ/ → /jomɯ/ 'read (non-past)'; /jom+sase/ → /jomase/ 'read (causative)'). These two cases would require conflicting rankings of MAX-IO and DEP-IO: epenthesis requires MAX-IO $\gg$ DEP-IO, while deletion requires instead DEP-IO $\gg$ MAX-IO. The solution Smith proposes, and which I adopt here, is to postulate a loanword-specific correspondence relation, the *SB-Correspondence* relation (for L2 *Source* and L1 *Borrowing*), and the corresponding set of faithfulness constraints: MAX-SB (segments in the source word must appear in the borrowing), DEP-SB (segments in the borrowing must appear in the source), and so on. Since non-loanwords satisfy these constraints by default (there being no L2 source word to be faithful to for these words), the ranking MAX-SB $\gg$ {DEP-SB, DEP-IO} $\gg$ MAX-IO allows for epenthesis repairs in loanword adaptation, but deletion repairs in verb morphology, as shown in Tables 2.2 and 2.3.

The second assumption which I will not be making in the analyses that follow is that all speakers share the same ranking for the constraints relevant to loanword adaptation. The existence of multiple possible adaptation strategies for the voiced geminate and coronal-/i/ patterns makes this assumption difficult to justify. In both of these cases, in addition to unnativized tokens of the loanwords in question, there are at least two

Table 2.2: Adaptation of *best*→[besɯto], showing loanword-specific epenthesis repairs (from Smith 2006)

	bɛst	CODACOND	MAX-SB	DEP-SB	DEP-IO	MAX-IO
a.	best	*!				
b. ☞	besɯto			**		
c.	be		*!*			

Table 2.3: Deletion repairs in verb morphology (from Smith 2006)

	jom+sase	CODACOND	MAX-SB	DEP-SB	DEP-IO	MAX-IO
a.	jomsase	*!				
b.	jom□sase				*!	
c. ☞	jomase					*

different repair strategies attested in nativized tokens from the 19th and early 20th centuries, which suggests that speakers in this time period had differing rankings for the relevant loanword faithfulness constraints. In other words, these speakers had differing preferences for preserving various features of the L2 source word. However, this is not a problem under Smith's (2006) proposal for SB-Correspondence constraints discussed above. Different speakers can have differing rankings for FAITH-SB constraints without affecting the rankings for the FAITH-IO and markedness constraints which govern the repair strategies seen in the non-loan phonology. As well, at the beginning stages of contact with English, it isn't surprising that different Japanese speakers would have had different rankings for SB constraints, since there were few loanwords existing in the language at that point and thus little evidence for deriving the "correct" ranking for these constraints.

2.3.2 Data collection

I obtained the data that follows on how the adaptation patterns were attested over time using `JMDICT` (Electronic Dictionary Research and Development Group 2003), a freely-available electronic dictionary of Japanese, and Arakawa (1977), a Japanese loanword dictionary. `JMDICT` was searched to generate a list of loanwords containing a particular non-native pattern, while Arakawa (1977) was used to find the approximate date of borrowing for each loanword in this list, and examples of how each loanword was attested in print over time. Each entry in Arakawa (1977) gives the source form from the language(s) the word was borrowed from, as well as several citations from newspapers, literary works, dictionaries, and other published material using the word from the entry. In general, Arakawa tries to give both the earliest attested uses, as well as more recent citations from the 1950's and 1960's if the word was still in use at the time of the dictionary's publication.

Since loanwords are generally written in *katakana* (Table 1.1 in Chapter 1), it is a relatively simple matter to determine what the intended pronunciation must have been for each cited form in Arakawa (1977). This is because the *katakana* writing system is orthographically shallow, in the sense that any valid sequence of *katakana* characters nearly always has only one possible pronunciation.[10] Distinctions such as palatalized or unpalatalized /t/ occurring before /i/ have a long history of being represented orthographically in *katakana* as well; for example, the ⟨ティ⟩ (/te/+/i/) spelling for unpalatalized [ti] is attested from at least the mid-19th century. It is likely, of course, that at least some historical Japanese speakers tended to read non-native orthographic sequences such as ⟨ティ⟩ in a nativized fashion (in this case, /ʧi/ instead of /ti/), since

[10]Going the other way, from pronunciation to orthography, there are only two cases in modern Japanese where ambiguity exists: /ʤi/, which may be spelled as either ⟨ジ⟩ (/ʃi/ + voicing mark) or ⟨ヂ⟩ (/ʧi/ + voicing mark); and /(d)zɯ/, which is represented using either ⟨ズ⟩ (/sɯ/ + voicing mark) or ⟨ヅ⟩ (/tsɯ/ + voicing mark). The ⟨ヂ⟩ and ⟨ズ⟩ characters are relatively rare in modern-day spellings, however.

earlier loanwords with source /ti/ are often spelled with either ⟨ティ⟩ /ti/ or ⟨チ⟩ /ʧi/. It has even been claimed, starting with Ichikawa (1930), that loanwords with orthographic voiced geminate stops, such as ベッド /beddo/ 'bed', were *always* pronounced /betto/, with a devoiced geminate. It is certain that nativization did take place for many loanwords, given that they are often attested in writing with a nativized spelling, especially for those words that are attested during the 19th century or earlier. Yet it can't be the case that *all* Japanese speakers *always* nativized foreign sequences like /ti/, while still making a distinction in writing between *katakana* pairs like ⟨ティ⟩ /ti/ and ⟨チ⟩ /ʧi/. For one thing, if orthographic pairs like ⟨ティ⟩ and ⟨チ⟩ were always pronounced the same, then we would expect words like チーズ /ʧiːzɯ/ 'cheese' and ティーズ /tiːzɯ/ 'tease' to be confused in writing, with /ʧiːzɯ/ sometimes being spelled as ティーズ instead of チーズ. Yet errors like this, where a native phonological sequence like /ʧi/ occurring in a loanword is written using a non-native orthographic string like ⟨ティ⟩, are never attested historically, to my knowledge. This kind of orthographic confusion does sometimes happen when modern-day Japanese speakers try to write an English word as they think it would be spelled using the Roman alphabet. For example, both /l/ and /ɹ/ in onset position in English are adapted into Japanese as /ɾ/, making homophones of such pairs as *flesh* and *fresh* (both adapted as /ɸɯɾeʃʃɯ/), and *flight* and *fright* (/ɸɯɾaito/). This leads to errors such as "flesh juice" being advertised on a restaurant's menu, or a "non-stop fright to Okinawa" on an airline poster.[11] The point is, when these kinds of spelling mistakes are made in both directions (such as source /l/ being spelled ⟨r⟩ and source /ɹ/ being spelled ⟨l⟩), then this indicates that the distinction between the two source sounds is neutralized for all speakers. When spelling mistakes only go in one direction, however (source /ti/ is sometimes spelled ⟨チ⟩, but source /ʧi/ is never spelled ⟨ティ⟩), then this suggests that the distinction between the two source sounds

[11]Photos of both of these examples can be found at the website <http://www.engrish.com/>. The name of the Japanese rock band Glay is a parody of these kinds of spelling mistakes, since it is an intentional misspelling of the word *gray* (Beech 1998).

must have been maintained by at least some speakers.

Another reason to believe that the attested spellings of loanwords reflects their degree of nativization is the fact that the attestation patterns of these non-native sequences can show sensitivity to both frequency and phonological neighborhood, as I will show with the data in section 2.3.5 on coronals before /i/. For instance, loanwords with source word-final /ti/ are the first to be attested with the ⟨ティ⟩ spelling (instead of ⟨チ⟩), followed by loans with word-initial /ti/ or /di/, then words with medial /ti/ or /di/. This difference in attestation times for these three sets of loanwords corresponds with their type frequency. Among all loans from source words containing /ti/, /ti/ occurs most frequently in word-final position, followed by word-initial position, then word-medial position. It is difficult to imagine how these spellings could be sensitive to the frequency and the phonological properties of the words involved, if they did not reflect the actual pronunciations used by at least some speakers at the time that these words were originally borrowed.

Since *katakana* is a phonologically-based orthography, I am only able to collect data on nativizations that involve pre-existing sounds in the native Japanese inventory (Figure 1.9 in Chapter 1), such as ⟨ティ⟩ /ti/ vs. ⟨チ⟩ /ʧi/, where /t/ is palatalized to /ʧ/ when occurring before /i/ in native Japanese words. This means that I am not able to quantify the degree of phonetic variation that existed in the production of non-native phonotactic sequences like /ti/ by historical speakers. This postulated phonetic variation, in many cases, would presumably have been an important factor affecting the final established form of the loanword (and is an important element of the model of loanword transmission that I develop in Chapter 5), but the effects of this variation can only be indirectly deduced, since only allophonic-level distinctions can be seen from the orthographic evidence I have collected here. In addition, there are some non-native patterns, such as unpalatalized /s/ before /i/, which have not even acquired a conventional representation in *katakana* until very recently. For example, [si] is generally spelled ⟨スィ⟩ (/sɯ/+/i/) in very recent loans, even though we would expect ⟨セィ⟩ (/se/+/i/) instead, on analogy with ⟨ティ⟩ (/te/+/i/). ⟨スィ⟩ is also sometimes used to represent loanwords with a source /swi/ or /swɪ/ sequence, such as スィッチ /sɯittʃi/ 'switch', and this may have something to do with its use for [si]. However, as far as I can tell words with ⟨スィ⟩ are unattested in Arakawa (1977), and loanwords with /si/ or /sɪ/ in the source word are always attested in this dictionary with ⟨シ⟩ /ʃi/. As well, there are some nonce spellings occasionally used in the 19th century, such as ⟨フヮ⟩ (/hɯ/+/wa/) used to represent the syllable [ɸa], or a small ⟨ラ⟩ used to indicate coda [r], which never became very common.

2.3.3 Palatalization of velars before /æ/

As was mentioned in section 2.2, velar stops occurring before /æ/ in English source words will sometimes be palatalized in the corresponding loanword, and sometimes not (compare キャット /k^{j}atto/ 'cat' with カタピラー /katapiraː/ 'caterpillar'). This palatalization process raises three questions:

1. Why does palatalization even occur in the first place, given that the velar in the English source word is not itself palatalized?
2. Why is it only the vowel /æ/ which triggers palatalization, and why do only velars palatalize in this environment?
3. Why is there variation in this adaptation pattern—in other words, why aren't all velars adapted as palatalized in the environment before /æ/?

The first two questions turn out to be related, and are due to the interaction between the articulatory phonetics of obstruents before front vowels in English, the adaptation of said vowels into Japanese, and phonotactic constraints on palatalization in Japanese. First of all, there is an allophonic variation in English between velars with a relatively front place of articulation, as in *key*, that occur before front vowels, and velars with a relatively back place of articulation, as in *coo* (Keating & Lahiri 1993). Fronted velars are thus likely to be perceived as palatalized by Japanese listeners, given their similarity in articulation (Akamatsu 1997).[12] This would predict that all velars occurring before front vowels should be adapted as palatalized velars into Japanese. However, only velars before /æ/ actually can be palatalized; the other possible source sequences here (/ki, kɪ, ke, kɛ/) are always adapted with a (phonologically) plain velar instead. This is because Japanese phonology does not make a palatalization contrast before front vowels (Ito & Mester 1995).[13] The /æ/ in /kæ/ however is usually adapted into Japanese as the non-front vowel [a], allowing palatalization to surface in this case. Note that in the few examples of /æ/ being adapted as [e] instead of [a], for example the early loan キャビン /kʲabiɴ/ 'cabin', which is often attested as ケビン /kebiɴ/, the velar is not indicated orthographically as being palatalized.

The question now is, why is velar palatalization before /æ/ not a regular process? To some extent, this seems to depend on when the word was first borrowed, judging from the date of the earliest attestation listed in the entry for each word in Arakawa (1977). A total of 431 tokens of 89 loanwords derived from source words containing

[12] Although I don't know of any perception studies that have actually tested this.

[13] The phonetic realization of non-coronal obstruents before front vowels in Japanese has changed over time. In the modern-day standard dialect, these obstruents are palatalized only before /i/ (Akamatsu 1997), but historically palatalization was triggered by both /i/ and /e/ (Shibatani 1999). Lange (1973: 35–37) argues on the basis of the spellings used in two historical texts, the *Ilopha (Iroha)* of 1492, a Korean textbook of Japanese, and the *Vocabvlario da Lingoa de Iapam*, a Japanese-Portuguese dictionary compiled by Jesuit missionaries in 1603, that palatalization on velars (and other obstruents, except for /s/) before /e/ was lost sometime in the 16th century. In any case, palatalization in this environment has always been a surface-level phonetic process; there has never been a phonological contrast between plain and palatalized non-coronal consonants before front vowels.

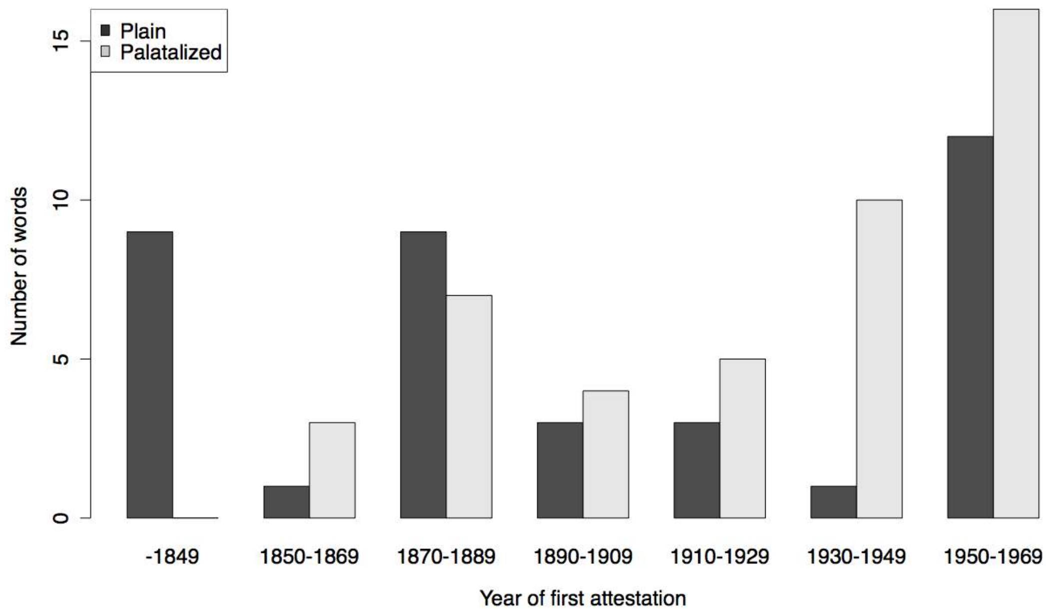


Figure 2.2: Adaptation patterns for velar stops before /æ/, by year of first attestation

a velar stop followed by /æ/ were collected from Arakawa (1977). The source words were also looked up in the Oxford English Dictionary (Pearsall & Trumble 2002) to ensure that they are pronounced with /æ/ in both RP and American dialects of English. As Figure 2.2 shows, there is some tendency for later borrowings to be palatalized: the 9 words in this data set that are first attested before 1850 have plain velars, while after 1890 the palatalized-velar loans outnumber the plain-velar loans borrowed in each time period, although even in the period 1950–1969 plain velars were still common in adaptations.

However, a better predictor of whether or not a given loanword will have a palatalized velar is the existence of similar-enough cognates in languages other than English, as Figure 2.3 shows. In this graph, a word is classified as "No cognates" if it is listed as having only an English source in Arakawa (1977), while it is classified as "Has cognates" if Arakawa lists multiple language sources for it (excluding Latin and Greek).

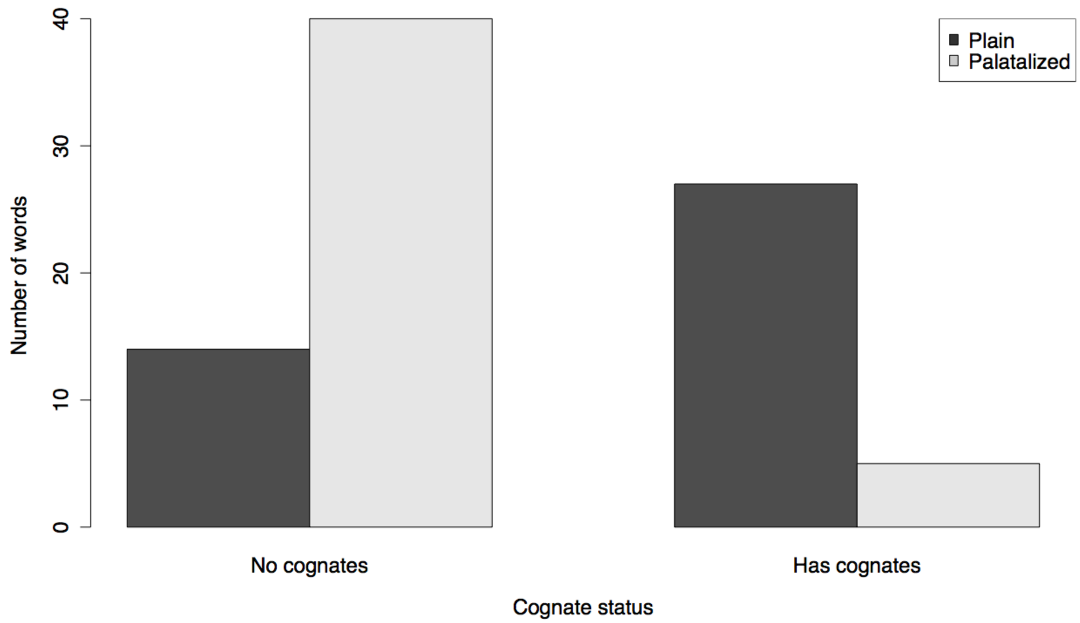


Figure 2.3: Adaptation patterns for velar stops before /æ/, by cognate status

About three-fourths of the words with no cognates turn out to have palatalized velars, while an even larger proportion of words with cognates have plain velars.

These two factors, date of first attestation and existence of cognates, together allow us to predict much of the variation in KÆ[14] adaptations. First, the few KÆ borrowings attested before about 1850, such as カテーテル /kateːterɯ/ 'catheter', ガス /gasɯ/ 'gas', and カタレプシー /katarepɯʃiː/ 'catalepsy', always have cognates in other languages (usually Dutch or German), and always have KA instead of KYA. In fact, given the contact situation at the time, these words, like other medical and scientific terms, are more likely to have been borrowed from Dutch or German directly than from English (Shibatani 1999: 148–9), in which case only the KA adaptation would be expected anyway, since there would have been no /æ/ present in the Dutch or German source

[14]From here on I will use the notation KÆ to refer to the source sequences /kæ/ and /gæ/ from English, and KYA and KA to refer to the corresponding palatalized (/k^ja/ and /g^ja/) and unpalatalized (/ka/ and /ga/) adaptations, respectively. KE will be used to refer to the (relatively rare) adaptation of KÆ as /ke/ or /ge/.

word.

Between 1850–1870, there start to be attested a few KÆ loans from English only, and these are adapted with KE at this time: ケビン /kebiɴ/ 'cabin', ケベージ /kebeːʤi/ 'cabbage', and ケップ /keppɯ/ 'cap'. There is also a single KYA adaptation attested during this time: キャプテン /kʲapɯteɴ/ 'captain', competing with the much earlier 17th century borrowing カピタン /kapitaɴ/ derived from the Portuguese word *capitão*. The other KÆ borrowings from this time have cognates in Dutch or German, and are adapted as KA: ガロップ /gaɾoppɯ/ 'gallop', カピトル /kapitoɾɯ/ 'capitol', カリフ /kaɾiɸɯ/ 'caliph', カシミア /kaʃimia/ 'cashmere', カセドラル /kasedoɾaɾɯ/ 'cathedral', ガロン /gaɾoɴ/ or ガルロン /gaɾɯɾoɴ/ 'gallon', カラクテル /kaɾakɯteɾɯ/ 'character', ガレリー /gaɾeɾiː/ 'gallery', カルシウム /kaɾɯʃiɯmɯ/ 'calcium'. The only exceptions are *caste*, which is initially attested during this time with KE (ケスト /kesɯto/), but afterwards is always attested with KA instead; and *cabinet*, which is also attested with KE (ケビネット /kebinetto/), but later shows up with KYA.

Moving on to the period from 1870–1890, new KÆ words that have cognates in German or French continue to be attested with KA only: カピタル /kapitaɾɯ/ 'capital', カンガルー /kaɴgaɾɯː/ 'kangaroo', カラバン /kaɾabaɴ/ 'caravan', カフェ /kaɸe/ 'café', カフェーン /kaɸeːɴ/ 'caffeine', カンタロープ /kaɴtaɾoːpɯ/ 'cantaloupe', カテゴリー /kategoɾiː/ 'category', カレンダー /kaɾeɴdaː/ 'calendar', カソリック /kasoɾikkɯ/ 'Catholic' (competing with the 18th century borrowing カトリック /katoɾikkɯ/), and カテキズム /katekizɯmɯ/ 'catechism' (competing with the 17th century borrowing カテキスモ /katekisɯmo/ from Portuguese *catechismo*). There are also many no-cognate loans dating from this time which are attested with KA as well: カリコ /kaɾiko/ 'calico', カンドルマス /kaɴdoɾɯmasɯ/ 'Candlemas', カンジータフト /kaɴʤiːtaɸɯto/ 'candytuft', カンブリック /kaɴbɯɾikkɯ/ 'cambric', カディー /kadiː/ 'caddy'. *Canvas* is attested both with KA (カンバス /kaɴbasɯ/) and KE (ケンバス /keɴbasɯ/). There are two other new no-cognate KE loans as well: ケッチャー /kettʃaː/ 'catcher' and ケンドル /keɴdorɯ/ 'candle' (compare with 'Candlemas'); and *cap*, which is first attested ca. 1867 with KE, shows up with KYA instead during the early 1870's (キャップ /kʲappɯ/). Finally, the word *character*, which was originally borrowed ca. 1850 as カラクテル /kaɾakɯteɾɯ/, is reborrowed during this time as ケラクター /keɾakɯtaː/ with the KE adaptation (note the coda /ɹ/ deletion here, which only occurs with loanwords from English), although it will later be attested with KA/KYA.

The variation in adaptations for no-cognate loans goes away after about 1890, and nearly all of these loans that are first attested after this date show the KYA adaptation only, for example in キャンデー /kʲaɴdeː/ 'candy', キャット /kʲatto/ 'cat', and キャッシュ /kʲaʃʃɯ/ 'cash', although there are a small number of words in this set of loans which show KA instead, such as カジュアル /kadʒɯaɾɯ/ 'casual'. Some of the earlier no-cognate loans which were originally borrowed with KA or KE are attested after 1890 with KYA instead: キャビン /kʲabiɴ/ 'cabin', キャリコ /kʲaɾiko/ 'calico', キャディー /kʲadiː/ 'caddy'. As for KÆ loans that do have cognates, these are still always attested with KA until about 1910, after which there are a few words which, although previously attested with KA, show up with KYA at this time instead, for example キャピタル /kʲapitaɾɯ/ 'capital', キャフェ /kʲaɸe/ 'café', and キャレンダー /kʲaɾeɴdaː/ 'calendar' (although of these three words, only /kʲapitaɾɯ/ seems to have replaced the earlier KA borrowing as the established form of the word). Finally, the word *guarantee*, despite having a transparent cognate in French (*garantie*), is attested only with the KYA adaptation (ギャランティ /ɡʲaɾaɴti/).

Table 2.4 summarizes the adaptation patterns for KÆ words by date and cognate status. I will now look at how to explain the changes in these adaptation patterns over

Table 2.4: Adaptation patterns for KÆ words, by date of first attestation and cognate status

Date attested	No cognates	Has cognates
before 1850	(no examples)	always KA
1850–1870	always KE (except kʲapɯteɴ)	usually KA (2 KE exceptions)
1870–1890	KA/KE (except kʲappɯ)	always KA
1890–1910	usually KYA	always KA
after 1910	usually KYA	usually KA, some KYA reborrowings

time, first by considering the role of on-line borrowing only, then by including the effects of transmission as well.

Adaptation of KÆ

As we saw above, among the entire set of KÆ words, there are three possible adaptations of KÆ: KA, KYA, and KE. For loans which come directly from English, then, it looks like there are two choices that a borrower needs to make: whether to adapt the non-native vowel /æ/ as either [a] or [e], and (if [a] was chosen) whether to palatalize the velar or not. Let us first consider the situation before 1890, when KÆ in no-cognate loans was generally adapted as either KA or KE; in this case, only the adaptation of the /æ/ is at issue here. The choice between KA and KE most likely stems from the low front vowel /æ/ being intermediate between the low back vowel /a/ and the mid front vowel /e/.[15] If we consider the distance in terms of features between these vowels, then a single change in the value of one of the features of /æ/ (+low, −back) can result in either /a/ (+low, +back) or /e/ (−low, −back). This change can take place either during perception, where the phonetic properties (specifically, the duration and F1 value) of the particular token of /æ/ being perceived will influence whether the borrower

[15] As noted in section 1.3, borrowing /æ/ directly seems not to have been an option, as there is no obvious way to represent it using the *kana* orthography.

Table 2.5: Adaptation of /kæ/→[ke], with IDENT-SB(back) ≫ IDENT-SB(low). The opposite ranking would result in /kæ/ being adapted as [ka] instead.

kæ	*æ	IDENT-SB(back)	IDENT-SB(low)
a. ka		*!	
b. kæ	*!		
c. ☞ ke			*

perceives it as either /a/ or /e/, or during the production stage, where the key factor here is the relative ranking of the constraints IDENT-SB(back) (preserve input values of [back]) and IDENT-SB(low) (preserve input values of [low]), as shown in Table 2.5. The variation between KA and KE could then be a result of different borrowers having different rankings for these two constraints. (I also assume here that there is a highly ranked constraint *æ, which prevents the non-native segment [æ] from surfacing in the output.)

After 1890, however, for the most part only the KYA adaptation is seen in new no-cognate KÆ loans. There are a couple of words (kʲapɯteɴ and kʲappɯ) that are attested before this date with KYA, and as I discussed before, the palatalization of the velar in these cases is probably an attempt to represent the perceptual frontness of English velars before front vowels like /æ/. Let us suppose that around 1890, borrowers generally started mapping source velars before /æ/ to palatalized velars,[16] which I assume are represented featurally as [−back] (and plain velars as [+back]). Then in this case, adapting KÆ as either KA or KE will entail an extra violation of IDENT-SB(back), and only KYA will survive as the most optimal candidate, no matter the relative rank-

[16]Perhaps on analogy with the already-existing loans kʲapɯteɴ and kʲappɯ, or perhaps due to influence from English teachers and instructional materials at this time. Note that there are a few examples of loanwords, like キャパシティ /kʲapaʃiti/ 'capacity' and キャリアー /kʲariaː/ 'career', which are sometimes attested with velar palatalization despite not having /æ/ in the source. This suggests that the change in mapping velars may have been, at least for some speakers, a kind of spelling pronunciation applied to any word with the initial sequence ⟨ca⟩ or ⟨ga⟩ (except for words where the ⟨a⟩ represents /eɪ/, like *cake* and *cage*).

Table 2.6: Adaptation of /kʲæ/→[kʲa]

	kʲæ	*Cʲe	*æ	IDENT-SB(back)	IDENT-SB(low)
a.	ka			**!	
b. ☞	kʲa			*	
c.	kæ		*!		
d.	kʲæ		*!		
e.	ke			*	*!
f.	kʲe	*!			*

ing of IDENT-SB(back) and IDENT-SB(low) (Table 2.6). Note that the other logically possible adaptation, KYE, never occurs because there is a high-ranking constraint *Cʲe forbidding palatalized consonants before /e/ (Ito & Mester 1995).

It is also possible to account for the no-cognate adaptation pattern using a change in constraint rankings only, by assuming that velars before /æ/ were always mapped to palatalized velars, even before 1890, and that there is an additional constraint *Cʲa (no palatalized consonants before /a/), analogous to *Cʲe. *Cʲa would have to be relatively low-ranked among markedness constraints, since /Cʲa/ sequences do occur in Yamato and Sino-Japanese words; in particular, it would be ranked below *Cʲe and *æ. Then KYA can be obtained with the ranking IDENT-SB(back) ≫ *Cʲa, while KA and KE can be obtained with the ranking *Cʲa ≫ IDENT-SB(back) (Table 2.7), with the choice between KA and KE being determined by the relative ranking of IDENT-SB(back) with IDENT-SB(low), as before.

While an analysis like the above relying solely on individual speaker adaptations can account for the range of variation in adaptation patterns seen in no-cognate loans, it turns out not to be possible to extend it to handle the loans with cognates as well. Of course, if a borrower is adapting a word like *category* directly from the French cognate

Table 2.7: Adaptation of /kʲæ/→[ke], with *Cʲa ≫ IDENT-SB(back). Ranking IDENT-SB(back) above *Cʲa would result in /kʲæ/ being adapted as [kʲa] instead.

	kʲæ	*Cʲe	*æ	*Cʲa	IDENT-SB(back)	IDENT-SB(low)
a.	ka				**!	
b.	kʲa			*!	*	
c.	kæ		*!		*	
d.	kʲæ		*!			
e. ☞	ke				*	*
f.	kʲe	*!				*

catégorie, then we would expect to see only the KA adaptation, since there is no /æ/ in the French source. The problem is how to represent, when a Japanese speaker is borrowing the English source, whether or not an English source word has cognates, since it is unreasonable to assume that all of the original borrowers of KÆ words with cognates must have taken them directly from French or German only and not from English. Postulating a constraint forbidding palatalized velars in source words with cognates would not be a very insightful solution, since that is simply restating the distribution in constraint form, and besides that, it would entail that all of the borrowers of these words had to have been in contact with French and German as well as English in order to know whether or not transparent cognates exist, which seems rather unlikely. A more plausible constraint to use here would be a constraint like *Cʲa proposed above, ranked above the relevant faithfulness constraints (IDENT-SB(back), in this case). This would correctly predict the KA adaptation[17] for a word like *category*, even if the English source is borrowed, but would also predict that no-cognate words like *cat* are always adapted with KA as well (Table 2.8), which is the wrong result. The problem here is that Gen will always generate both plain and palatalized versions of all loanwords, regardless of

[17]Actually, this is only true if IDENT-SB(low) ≫ IDENT-SB(back); the reverse ranking would predict KE as the outcome for all KÆ words.

Table 2.8: Incorrect adaptation of *cat*→[katto] instead of [kʲatto]

kʲæt	*Cʲe	*Cʲa	IDENT-SB(low)	IDENT-SB(back)
a. kʲatto		*!		*
b. ☹ katto				**

cognate status, and any ranking of constraints which allows KA to win out over KYA in cognate loans will thus allow KA to win out over KYA in no-cognate loans too.

Transmission of KÆ loans

Since it is not possible to derive the correct distribution of KA and KYA adaptations from the effects of on-line borrowing only, I will now consider how the transmission of KÆ loans could have resulted in the KYA adaptation being dispreferred in loans with transparent cognates. I will assume here that borrowers are not generally aware of the cognate status of the words they are borrowing, but instead simply adapt the loanword based on a single source language. Then, no-cognate loans like *cat* will (after 1890) be borrowed with KYA, while cognate loans like *category* will be borrowed by some speakers with KYA, basing it on the English source, and by other speakers with KA instead, basing it on either the French (*catégorie*) or German (*Kategorie*) source. This means that, for cognate loans, there will be both KYA and KA variants spreading through the speech community, while for no-cognate loans, only the KYA variant will exist. While for transparent cognates like *category*, *catégorie*, and *Kategorie*, the only difference between the English-source and other-source variants will be the palatalization of the velar, for less similar cognates, such as *captain* and *Kapitän*, there will be other differences as well, some examples of which are given in Table 2.9. For example, in the pair /kʲapɯteɴ/ and /kapitaɴ/, the English-source variant has an epenthetic /ɯ/ breaking up the /pt/ cluster in the English source, whereas the corresponding vowel in the other-source variant is an [i], reflecting the /i/ present in the non-English sources. Likewise, in /kʲaɾakɯtaː/ and /kaɾakɯteɾɯ/, the first member of the pair shows coda /r/ vocalization, which is what is generally seen in borrowings from English, while the second preserves the coda /r/ from the source word by epenthesizing /ɯ/, which occurs in borrowings from non-English sources (for example, アムール /amɯːɾɯ/ 'amour', from French). Although in some doublets, neither member shows velar palatalization, note that, when velar palatalization does occur, it is nearly always the English-source member of the pair that displays it.[18]

Given that in many cases of less similar cognate pairs, both KYA and KA adaptations are attested, it seems likely then that KYA and KA variants of transparent cognate pairs can co-occur as well. Which variants a particular speaker is exposed to would then depend on their position in the social network relative to the borrowers of the different variants. Those speakers who are relatively close to one of the KYA borrowers, in terms of network distance (the minimum number of links required to travel from the borrower to the speaker in the network), and are relatively far from the KA borrowers, will mostly be exposed to the KYA variant, and presumably would use it in their own speech (thus furthering its spread). Those speakers who are near one of the KA borrowers but far away from the KYA borrowers, however, will end up using the KA variant instead. The interesting case, then, is of those speakers who are positioned in the network such that they are likely to be exposed to both KYA and KA variants, and thus will have to make a choice between using one or the other when they try to produce the loanword themselves. Since (with few exceptions) cognate loans are always attested with KA, it

[18]The one exception is the borrowing of *cabinet* from French, which is attested three times in the entry in Arakawa (1977) as /kabine/, and once as /kʲabine/. However, the latter may actually be a truncated form of the English borrowing /kʲabinetto/, rather than a borrowing directly from the French source word. (Although most loanword truncations are only two moras long, there are a fair number which are three moras, such as /teɾebi/ < /teɾebiʤoɴ/ 'television', so this is not entirely implausible.) If so, then the existence of /kʲabine/ would not be a true exception to the generalization that velar-palatalized forms are always derived from the English member of the cognate set.

Table 2.9: Palatalization in loanword doublets. Fr=French; De=German; Nl=Dutch; Pt=Portuguese; Ru=Russian. The dates given are of the citations listed in Arakawa (1977) for each word.

English source	Loanword	Dates attested	Non-English source	Loanword	Dates attested
captain	kʲapɯteɴ	1850s	*capitão* (Pt)/*Kapitän* (De)/ *kapitan* (Ru)	kapitaɴ	17th–18th c., 1850s
			kapitein (Nl)/*Kapitän* (De)	kapite(ː)ɴ	1750s–1890s
cabinet	kʲabinetto	1860s, 1930s, 1950s	*cabinet* (Fr)	k$^{(j)}$abine	1890s
cashmere	kaʃimia	1860s, 1900s, 1950s	*cachemire* (Fr)/*Kaschmir* (De)	kaʃimiːɾɯ	1870s–1910s
catholic	kasoɾikkɯ	1870s–1910s, 1950s	*katholiek* (Nl)/*catholique* (Fr)/ *Katholik* (De)	katoɾikkɯ	18th c.
character	kʲaɾakɯtaː	1870s–1900s, 1940s	*karakter* (Nl)/*caractère* (Fr)/ *Charakter* (De)	kaɾakɯteɾɯ	1850s–1860s, 1900s
caricature	kaɾikatʃɯa	1900s, 1950s	*caricature* (Fr)/*Karikatur* (De)	kaɾikatʃɯːɾɯ	1900s–1910s
	kaɾɯk(a/e)tʃɯa	1910s–1920s, 1950s			
gasoline	gʲasɯɾiɴ	1900s, 1950s	*Gasolin* (De)	gasoɾiɴ	1910s–1920s

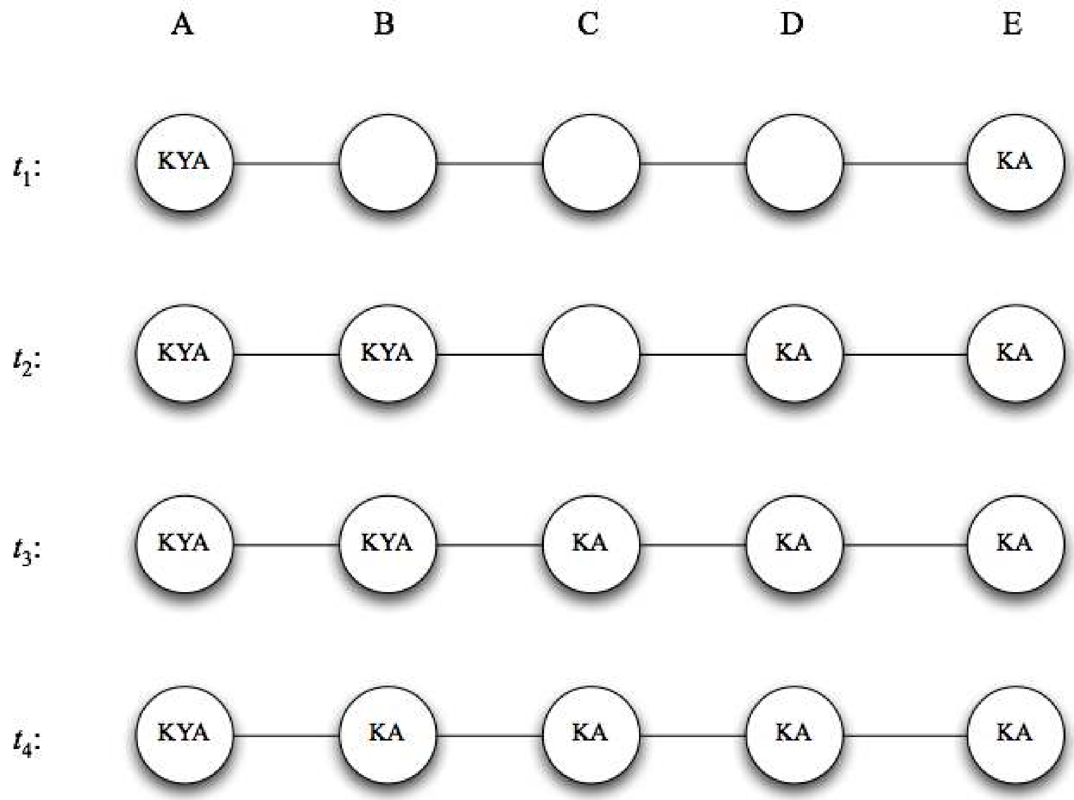


Figure 2.4: KYA and KA variants of a loanword spreading in a simplified social network.

seems that these speakers who are exposed to both variants end up choosing KA over KYA, and this somehow results in KA eventually becoming the established form of the loanword.

To see how this happens, consider the transmission of KYA and KA variants in a highly simplified social network consisting of five speakers arranged in a line, so that each speaker only communicates directly with the speakers to the left and right of her in the line (Figure 2.4). At time t_1, the two speakers A and E at the opposite ends of the line borrow the KYA and KA variant, respectively. We then expect speaker B to learn the KYA variant from speaker A, and likewise speaker D to learn the KA variant from speaker E, as shown at time t_2. Now speaker C will be exposed to both the KYA variant (from speaker B) and the KA variant (from speaker D). Assuming that, whenever a speaker is exposed to both variants, she will always use the KA variant in her own speech, then speaker C will end up learning KA, as shown at time t_3. Speaker B is now being exposed to the KA variant (from speaker C), in addition to the KYA variant (from speaker A), resulting in B switching to the KA form at time t_4.[19] In fact, no matter how many speakers there are, if they are allowed to switch which variant they use at any time, then what will happen is that, if there is at least one borrower who introduces the KA variant into the network, all speakers will eventually switch to the KA variant[20] as it spreads through the network.

Why then would a speaker, faced with variation in a loanword between KYA and KA, always end up choosing KA? KYA does not seem to be phonologically marked in Japanese, since examples of palatalized obstruents followed by /a/ can be found in all strata (McCawley 1968). One possibility is the social prestige of speakers in specialist communities (like medicine, arts and literature, and so on) who had some knowledge of French or German. Given their language ability, these speakers would have carried some clout with their monolingual peers. These high-prestige speakers, upon being exposed to a new KÆ word, would most likely have favored the KA pronunciation over KYA if they happened to know the French or German cognate, believing this pronunciation to be more "correct" than the English-based KYA variant, and perhaps would even have tried to correct other speakers in their immediate circle who used KYA instead. This would lead to KA becoming the established form within the specialist community, and ultimately within the broader Japanese speech community, if the loanword came to be used outside of specialist circles.

[19]This is assuming that speaker B is allowed to switch to KA on the basis of new examples of the loanword, even though KYA was the form originally learned. Whether this is possible would depend on the details of the learning algorithm that speakers use to acquire new words.

[20]Except possibly the KYA borrowers, who are also influenced by external evidence from the English source of the loanword.

Table 2.10: Adaptation of *café*→[kaɸe]

	…	*[−back]	*g	*k	…
a. ☞ kaɸe				*	
b. kʲaɸe		*!		*	

Another possible reason for favoring KA over KYA is that palatalized consonants are articulatorily more complex than plain consonants, making KA easier to produce than KYA, all else being equal. This can be formalized in OT using a low-ranked constraint *[−back] (Kager 1999: 127)[21] which is ranked below faithfulness constraints like IDENT-IO(back) and IDENT-SB(back). This ranking will ensure that palatalization in the input will be preserved in the output, as in native words with an underlying palatalized segment. I assume that loanwords like *cat* are also represented underlyingly with a palatalized /k/, since only the KYA variant is attested, and so presumably Japanese speakers are only exposed to the KYA variant and thus would postulate the UR as /kʲatto/ instead of /katto/ (the latter of which they would have no evidence for from other speakers' productions of the word). When words like *café* were borrowed, however, Japanese speakers were probably exposed to both the KYA and KA variants, as I argued above, and ended up selecting the KA variant. Suppose that, when a speaker is faced with conflicting evidence like this for the UR of a new word being learned, and the variants differ by only one feature (in this case, the value of the feature [back] for the initial /k/), they always assume that the UR is the least marked variant according to the ranking of the markedness constraints in their grammar, as in Table 2.10. While both forms violate various low-ranked markedness constraints like *k, only the KYA variant violates *[−back], and so the KA variant is instead picked as the UR for the loanword.

[21] Kager (1999) proposes this constraint, along with *[+low] and *[+round], to explain why epenthetic vowels tend to be centralized and non-low, cross-linguistically. Since I am assuming palatalization is represented using the [back] feature, then *[−back] also has the effect of ruling out palatalized consonants when a plain counterpart is available.

Thus, in the process of loanword transmission, no-cognate loans will always show the KYA adaptation, while cognate loans will end up as KA instead (even if some borrowers introduced KYA).[22]

2.3.4 Gemination of word-final voiced obstruents

I now turn to my next example of loanword adaptation, namely, the occurrence of voiced geminates in Japanese loanwords. Even though none of the main source languages for loanwords (English, French, and German) make a distinction between singleton and geminate consonants, geminates do occur in many loans from these languages, for example in ペット /petto/ 'pet' and エッセイ /esseː/ 'essay'.[23] Here I will focus on the case of word-final gemination in English source words, since it is the easiest to state the conditions under which it occurs. In this set of loans, a word-final voiceless stop or affricate will always be adapted as a geminate if it follows a lax vowel (Tsuchida 1995). So in addition to *pet* above, we also have examples like ポケット /poketto/ 'pocket', アップ /appɯ/ 'up', チェック /ʧekkɯ/ 'check', and タッチ /tatʧi/ 'touch'.[24]

Tsuchida (1995) gives the most complete OT account of gemination in Japanese loanwords. Her analysis of word-final gemination involves a constraint ALIGN (which I will call ALIGN-SB, since it is a type of SB-correspondence), which requires the right edge of the source word to align with the right edge of a syllable in the loanword.

[22] Of course, it is not the mere existence of a cognate in another language which blocks velar palatalization; the cognate also has to be borrowed by one or more speakers and be circulating through the speech community in order for it to win out over the form borrowed from English.

[23] Geminates can also occur in loanwords from Italian, such as スパゲッティ /sɯpagetti/ 'spaghetti', which have a geminate in the source word.

[24] This is (usually) true of the voiceless fricative /ʃ/ as well, as in フィッシュ /ɸiʃʃɯ/ 'fish'. However, words with final /f/ or /s/ in English generally fail to geminate in this environment; the only exceptions listed in `JMDICT` are スタッフ /sɯtaɸɸɯ/ 'staff; stuff', マ(ッ)ス /ma(s)sɯ/ 'mass', and キ(ッ)ス /ki(s)sɯ/ 'kiss'. Note that for the last two forms, the ungeminated form is actually the more common one.

Table 2.11: Adaptation of *pet*→[petto]

pɛt	*CODAPLACE	MAX-SB	ALIGN-SB	DEP-SB
a. pet	*!			
b. pe		*!	*	
c. pe.to			*!	*
d. ☞ pet.to				**

Table 2.12: Adaptation of *date*→[deːto], *[deːtto]

deɪt	*CODAPLACE	*3μ	MAX-SB	ALIGN-SB	DEP-SB
a. deːt	*!				
b. deː			*!	*	
c. ☞ deː.to				*	*
d. deːt.to		*!			**
e. det.to			*!		**

This constraint is dominated by *CODAPLACE (codas are limited to placeless nasals and first segments of geminates; analogous to CODACOND from Ito & Mester (1995) discussed in section 1.3) and LOANWORD CORRESPONDENCE (preserve segments of source word; equivalent to MAX-SB), forbidding deletion of the coda to take place; and itself dominates FILL (no epenthesis; equivalent to DEP-SB), allowing gemination to occur in the resulting loanword, as shown in Table 2.11.

Gemination only takes place after lax vowels, which are adapted into Japanese as short vowels. Tense vowels, which are adapted as long vowels, block gemination because the resulting trimoraic syllable would violate another one of Tsuchida's (1995) constraints, *SUPERHEAVY SYLLABLE (which I will abbreviate here as *3μ), as shown in Table 2.12. Lovins (1975: 84) also gives an explanation for gemination in loanwords which is very similar to Tsuchida's account, namely that it occurs as an attempt to pre-

Table 2.13: Gemination and epenthesis alternations in Sino-Japanese compounding (examples from Katayama 1998: 128–9)

dat + kai	→ da<u>kk</u>ai	*datsɯkai	'resign'
tot + ʃiɴ	→ to<u>ʃʃ</u>iɴ	*totsɯʃiɴ	'dash'
hat + keɴ	→ ha<u>kk</u>eɴ	*hatsɯkeɴ	'discover'
hat + teɴ	→ ha<u>tt</u>eɴ	*hatsɯteɴ	'development'
dat + boː	→ *dabboː	da<u>tsɯb</u>oː	'take off one's hat'
tot + geki	→ *toggeki	to<u>tsɯg</u>eki	'charge'
hat + geɴ	→ *haggeɴ	ha<u>tsɯg</u>eɴ	'utterance'
hat + deɴ	→ *haddeɴ	ha<u>tsɯd</u>eɴ	'generation of electricity'

serve the closed nature of the final syllable of the source word, but is blocked when the resulting syllable would be trimoraic.

The interesting case is what happens when the word-final obstruent is voiced. Voiced geminates never occur underlyingly in Yamato, Sino-Japanese, or Mimetic words,[25] and there are a few cases of alternations where voiceless geminates alternate with either /tɯ/ (=[tsɯ]) or /ɴ/ followed by the corresponding voiced singleton, suggesting an active constraint against voiced geminates. For example, a root-final /−t/ causes gemination of a following voiceless consonant in Sino-Japanese compounds, but surfaces as [tsɯ] (/t/ + an epenthetic vowel) before voiced consonants (Table 2.13). In Mimetic roots of the form /(C)VXC_1V−ɾi/, X=C_1 when C_1 is voiceless, but X=/ɴ/ instead when C_1 is voiced (Table 2.14),[26] and a similar alternation between voiceless geminates and nasal—voiced obstruent clusters is seen in the past tense forms of verbs (Table 2.15).[27]

[25] Although Akamatsu (1997) notes some cases of native words, such as /kɯdaɾanai/ 'absurd' or /sɯgoi/ 'terrific', where the voiced stop in the word may be produced with a lengthened stop closure for the purposes of emphasis. However, this would be a surface-level phonetic process only; there is no reason to believe that the underlying forms of these words would have geminate /dd/ or /gg/.

[26] This is a different analysis from Ito & Mester (1995), who presumably consider the voiceless geminate forms to be a violation of *NT, the constraint against voiceless obstruents appearing after /ɴ/. Although they are not explicit about how mimetic forms with /−ɾi/ would be represented, it seems that a form like [pattaɾi] 'pitter-patter' would underlyingly be /paɴtaɾi/ under their analysis, with the underlying /ɴ/ assimilating in place to the following /t/, and also having its [nasal] feature delinked to avoid violating *NT.

[27] Although in this case, the only clear example comes from verb roots ending in /−b/. Roots ending in /−d/, /−ʤ/, or /−z/ do not happen to exist in the language, while the past tense forms of roots ending in /−g/ reflect a historical lenition of velars at morpheme boundaries where the velar changes to

Table 2.14: Gemination and nasal-obstruent alternations in Mimetic roots (examples from Katayama 1998: 129)

pattaɾi	'pitter-patter'
nikkoɾi	'smiling'
ʃiɴnaɾi	'supple'
ʃiɴmiɾi	'solemn; serious'
ʃoɴboɾi	'downhearted'
uɴzaɾi	'tedious; boring'

Table 2.15: Gemination and nasal-obstruent alternations in past tense forms

mat−	+ −ta	→	matta	'waited'
kiɾ−	+ −ta	→	kitta	'cut (past)'
kaw−	+ −ta	→	katta	'bought'
asob−	+ −ta	→	asoɴda	'played'
jom−	+ −ta	→	joɴda	'read (past)'

Thus, we might expect voiced obstruents from English words in the gemination environment (word-finally after a lax vowel) not to surface as geminates, but instead be repaired in some way in the resulting loanword. The traditional claim is that this is indeed what happens, and that illegal voiced geminates are repaired via devoicing: Ichikawa (1930) gives the examples キット /kitto/ 'kid', ベット /betto/ 'bed', ヘットライト /hettoɾaito/ 'headlight', and オペラパック /opeɾapakkɯ/ 'opera-bag',[28] and these examples have been widely cited in later discussions of voiced geminates, for example in Lovins (1975). A possible OT analysis of these facts would involve a constraint *DD (no voiced geminate obstruents; Ito & Mester 1995) being ranked above the faithfulness constraint IDENT-SB(voi) (preserve values of [voi] feature of source word segments), as in Table 2.16.[29] Hayes & Steriade (2004), following Ohala (1983), suggest that *DD

[i] (Shibatani 1999), for example in /ojog+ɯ/ 'to swim', /ojog+da/ → [ojoi̱da] 'swam'.

[28]Note also in this form the (unexpected) devoicing of the /b/ from *bag*.

[29]Tsuchida (1995) proposes two constraints which would correspond to Ito & Mester's (1995) *DD, one which rules out gemination of voiced fricatives and /b/ (which Tsuchida analyzes as [+cont]), and the other which rules out gemination of voiced stops and affricates, with the ranking *VOICED [+CONT] GEMINATE ≫ ALIGN-SB ≫ *VOICED [−CONT] GEMINATE ≫ FILL. The ranking of these two constraints is used to explain why /zz/ and /bb/ sequences are so rare in loanwords compared to other voiced geminates.

Table 2.16: Adaptation of *bed*→[betto], with *DD ≫ IDENT-SB(voi)

bɛd	MAX-SB	ALIGN-SB	*DD	IDENT-SB(voi)	DEP-SB
a. be	*!	*			
b. bedo		*!			*
c. beddo			*!		**
d. ☞ betto				*	**

has a phonetic basis, namely, that it is more difficult to maintain voicing in obstruents than in other segments because voicing requires continuous airflow through the glottis, and in obstruents by definition the airstream is blocked.[30]

Lovins (1975) also shows that there is some variation in this set of loans as to whether the geminate is voiced or not. She gives the forms ベッド /beddo/ and バッグ /baggɯ/ as other possible forms for *bed* and *bag*, and in addition, she lists many other loanwords which are generally attested with a voiced geminate only, such as スキャッブ /sɯkʲabbɯ/ 'scab', スノッブ /sɯnobbɯ/ 'snob', グリッド /gɯriddo/ 'grid', ロッド /roddo/ 'rod', バッジ /badʤi/ 'badge', ジャッジ /ʤadʤi/ 'judge', ビッグ /biggɯ/ 'big', ドッグ /doggɯ/ 'dog', and エッグノッグ /eggɯnoggɯ/ 'eggnog', as well as several which are attested instead with a voiced singleton (which is another possible strategy for avoiding a voiced geminate), such as クラブ /kɯrabɯ/ 'club', プラグ /pɯragɯ/ 'plug', ナツメグ /natsɯmegɯ/ 'nutmeg', and カタログ /katarogɯ/ 'catalog'. Lovins also notes that there are some examples of doublets derived from the same source word which show variation in whether the final stop is geminated or not, for ex-

[30]This historical pattern of devoicing voiced geminates seems to have a modern-day counterpart. Kawahara (2006), citing corpus research by Nishimura (2003), shows that devoicing of voiced geminates may optionally occur in words where they appear with another voiced obstruent, such as [beddo] or [betto] 'bed', and [guddo] or [gutto] 'good'. Devoicing does not occur when the geminate occurs with a voiceless obstruent or with a sonorant, as in [kiddo], *[kitto] 'kid', and [webbɯ], *[weppɯ] 'web'. In Chapter 4 I discuss Kawahara's (2006) analysis of this pattern as a conflict between two faithfulness constraints governing voicing in singletons and geminates, as well as Pater's (2008) alternative analysis using cumulative constraint interaction in Harmonic Grammar.

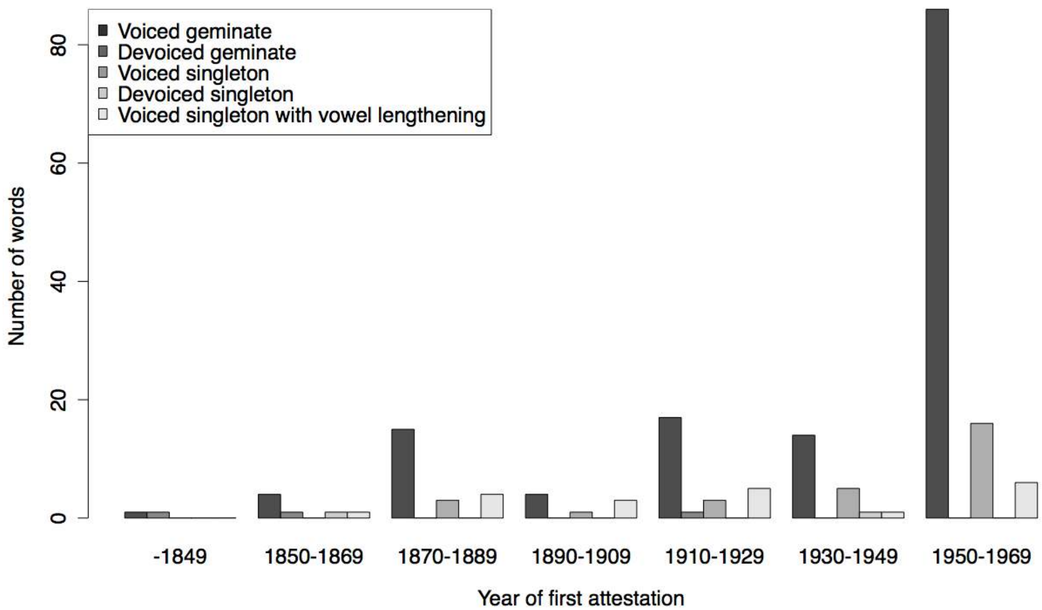


Figure 2.5: Adaptation patterns of word-final voiced obstruents, by date of first attestation

ample ギャッグ /gʲaggɯ/ or ギャグ /gʲagɯ/ 'gag', ロッブ /ɾobbɯ/ or ロブ /ɾobɯ/ 'lob (tennis shot)', and ノッブ /nobbɯ/ or ノブ /nobɯ/ 'knob'.

The data I have collected from Arakawa (1977) also suggest that the situation is far more complex than a simple adaptation rule devoicing voiced geminates in loanwords. Figure 2.5 shows the frequency of various adaptation patterns for 205 source words with a word-final voiced obstruent following a lax vowel. It is clear from this graph that starting in 1870, voiced geminates are attested more frequently in each time period than all other adaptation patterns combined. The next most common adaptation patterns are voiced singletons (as in クラブ /kɯɾabɯ/ 'club') and voiced singletons with lengthening of the preceding vowel (as in イメージ /imeːʤi/ 'image'). Voiceless geminates, however, are rarely attested in any time period. Among the entire set of words collected from Arakawa, the only ones which are listed with three or more attestations with a

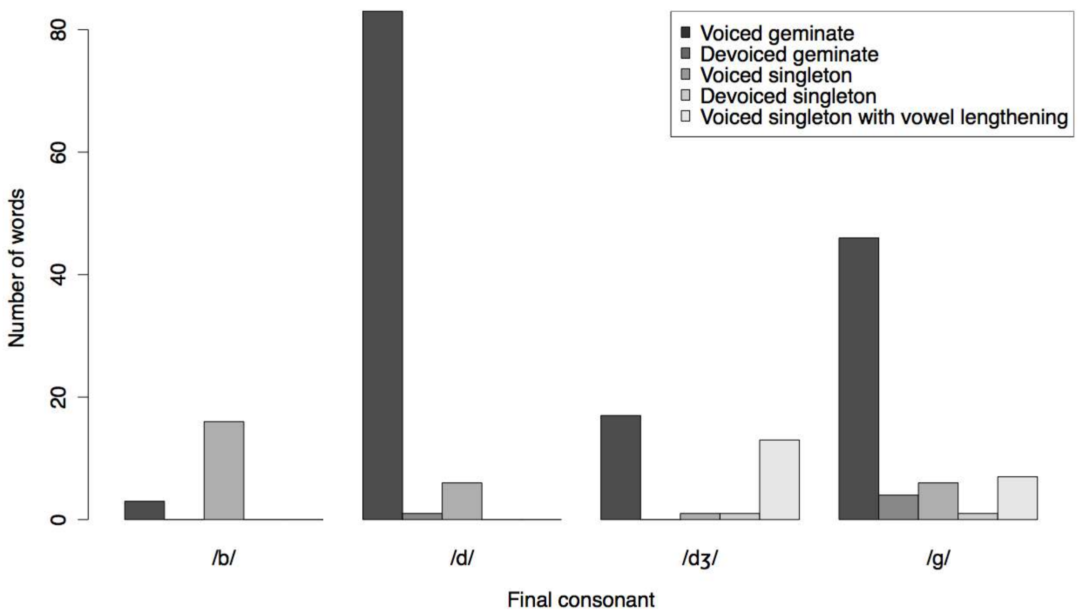


Figure 2.6: Adaptation patterns of word-final voiced obstruents, by final consonant. /ʤ/ and /g/ data include words spelled ⟨−age⟩ or ⟨−ogue⟩.

voiceless geminate are ベット /betto/ 'bed',[31] ハンドバック /haɴdobakkɯ/ 'handbag', オペラパック /operapakkɯ/ 'opera-bag', and ブルドック /bɯrɯdokkɯ/ 'bulldog'. Of these, both *bed* and *opera-bag* are also attested with alternate forms, ベッド /beddo/ and オペラバッグ /operabaggɯ/, that have a voiced geminate.

A striking pattern emerges if we compare adaptation patterns by the place of articulation of the final consonant of the source word, as in Figure 2.6. Here we can see that final /d/ or /g/ is nearly always adapted as a voiced geminate, final /ʤ/ is about equally likely to be adapted as a voiced geminate or a voiced singleton with vowel lengthening, and final /b/ is usually adapted as a voiced singleton. Note that the vowel lengthening pattern applies only to words spelled with final ⟨−age⟩ or ⟨−ogue⟩, such as パッケージ /pakkeːʤi/ 'package', ソーセージ /soːseːʤi/ 'sausage', and プロローグ /pɯroroːgɯ/

[31] Arakawa notes that *bed* has cognates in German (*Bett*) and Dutch *bed*, both of which are pronounced with a final [t], which may have been another source for the voiceless geminate form /betto/.

Table 2.17: Adaptation patterns for voiced geminates, by date of first attestation and final consonant. VC=voiced singleton; VC̥=voiceless singleton; VCC=voiced geminate; VC̥C̥=voiceless geminate; VVC=voiced singleton with vowel lengthening. /ʤ/ and /g/ columns exclude words spelled ⟨−age⟩ or ⟨−ogue⟩.

Date attested	/b/	/d/	/ʤ/	/g/
before 1870	(no examples)	2 tokens VCC, 3 tokens VC̥C̥	(no examples)	2 tokens VC̥, 2 tokens VCC
1870–1910	1 VC (/kɯrabɯ/ ‘club’), 2 VCC (/mobbɯ/ ‘mob’, /sɯnobbɯ/ ‘snob’)	VCC, sporadic VC/VVC	VCC, 1 token VVC	2 tokens VC, 5 tokens VCC
1910–1950	VC (except /basɯtabbɯ/ ‘bathtub’)	VCC, 1 token VVC	4 tokens VCC, 3 tokens VC	VCC, sporadic VC/VVC/VC̥C̥
after 1950	VC (most are compounds with /kɯrabɯ/)	VCC, sporadic VC	VCC	VCC, sporadic VC

‘prologue’. This seems to be a spelling pronunciation based on monosyllabic words with ⟨−age⟩ or ⟨−ogue⟩, such as ステージ /sɯteːʤi/ ‘stage’, ページ /peːʤi/ ‘page’, or ボーグ /boːgɯ/ ‘vogue’. Gemination does not occur in these words because the original borrowers apparently misread them as having a tense vowel followed by a coda consonant (i.e. /pækeɪʤ/ instead of /pækəʤ/ for *package*), causing the vowel to be adapted as long instead of short, and gemination in this environment (*pakkeːdʤi) would violate the *3μ constraint from before.

The remaining adaptation patterns interact with the time of borrowing of the loans, as shown in Table 2.17, although note that in all time periods, there can be found examples of loanwords attested with voiced geminates. In general, there seems to be greater variation in adaptation patterns possible before 1890, but after this date final /d/ and /ʤ/ are both adapted as voiced geminates, with only sporadic instances of singletons or singletons with vowel lengthening occurring, while final /g/ shows somewhat more variation until about 1950, after which it is more consistently adapted as a voiced geminate. Final /b/, on the other hand, is almost always adapted as a singleton in all time

Table 2.18: Adaptation of *bed*→[beddo], with {ALIGN-SB, IDENT-SB(voi)} $\gg$ *DD

	bɛd	ALIGN-SB	IDENT-SB(voi)	*DD	DEP-SB
a.	bedo	*!			*
b. ☞	beddo			*	**
c.	betto		*!		**

periods. Note also that devoicing, while sometimes possible before 1870, almost never occurs in loanwords first attested after this date.

Adaptation of voiced geminates

Leaving aside the vowel-lengthening cases, which are for the most part spelling pronunciations as I mentioned above, there are three possible adaptations for word-final voiced stops following a lax vowel: VCC (voiced geminate), VC̥C̥ (voiceless geminate), and VC (voiced singleton). Tsuchida's (1995) analysis shows that, for modern-day Japanese speakers, {CODACOND, *3μ, MAX-SB} $\gg$ ALIGN-SB $\gg$ *DD $\gg$ DEP-SB, resulting in gemination of word-final stops after a lax vowel, even if the stop is voiced. However, as I discussed in section 2.3.1, for 19$^{\text{th}}$-century speakers, the rankings of FAITH-SB constraints with respect to each other and with respect to native constraints were not consistent from speaker to speaker. In particular, the ranking of the constraint IDENT-SB(voi), which preserves the voicing of source word segments, with respect to *DD determines whether a voiced stop in the gemination environment is borrowed as a voiced or voiceless geminate. Assuming ALIGN-SB $\gg$ DEP-SB, if ALIGN-SB and IDENT-SB(voi) $\gg$ *DD, the geminate will be voiced (Table 2.18); if IDENT-SB(voi) $\ll$ *DD, on the other hand, the geminate will be voiceless (Table 2.16). The VC adaptation pattern, meanwhile, can be derived either by ranking *DD and IDENT-SB(voi) above ALIGN-SB, as in Table 2.19, or by ranking DEP-SB above ALIGN-SB, in which case both VCC and

Table 2.19: Adaptation of *pub*→[pabɯ], with {*DD, IDENT-SB(voi)} ≫ ALIGN-SB

pʌb	IDENT-SB(voi)	*DD	ALIGN-SB	DEP-SB
a. ☞ pabɯ			*	*
b. pabbɯ		*!		**
c. pappɯ	*!			**

VC̥C̥ will lose to VC because they each entail an extra violation of DEP-SB. There are no other possible rankings of DEP-SB, ALIGN-SB, and IDENT-SB(voi) with respect to *DD which result in an adaptation that is not either VCC, VC̥C̥, or VC.[32]

Yet it is more difficult to explain why the use of the VCC/VC̥C̥/VC adaptation strategies depend on the place of articulation of the word-final obstruent. As I noted above, final /d/ and /ʤ/ are generally adapted as VCC, final /b/ as VC, and final /g/ shows a tendency towards VCC, but with a great deal of variation until recently. Tsuchida (1995), in attempting to explain the synchronic adaptation pattern, suggests that the *DD constraint should be split into two constraints, one which rules out /bb/ (which I will call *BB) and the other which rules out the other possible voiced geminates,[33] with *BB ≫ *DD, as in Table 2.20. (It is also necessary for IDENT-SB(voi) ≫ ALIGN-SB, or else the candidate [pappɯ] would win out over [pabɯ]. To my knowledge, VC̥C̥ is never attested as an adaptation for word-final /b/ in Arakawa (1977).)

This is not a particularly insightful analysis, of course. It merely stipulates that /bb/ is phonologically more marked than /dd/, /dʤ/, and /gg/, but without explaining why.

[32]Note that the VVC adaptation is harmonically bound by the other three adaptations with respect to the constraints being considered here. VVC always loses to VC on violations of DEP-SB, and to VCC and VC̥C̥ on ALIGN-SB.

[33]Actually, what Tsuchida (1995) does is postulate a constraint *VOICED [+cont] GEMINATES ruling out voiced geminate fricatives, and then argues that this constraint applies to /b/ too, since /b/ is variably lenited to [β] intervocalically in Japanese. This analysis seems doubtful to me, though, since /b/ clearly patterns phonologically with the other stops in Japanese, and not with fricatives, making it unlikely that it is represented underlyingly as [+cont].

Table 2.20: Adaptation of *pub*→[pabɯ], with *BB ≫ IDENT-SB(voi) ≫ ALIGN-SB ≫ *DD

pʌb	*BB	IDENT-SB(voi)	ALIGN-SB	*DD	DEP-SB
a. ☞ pabɯ			*		*
b. pabbɯ	*!				**
c. pappɯ		*!			**

Lovins (1975) and Katayama (1998) both discuss possible phonetic reasons for why gemination is more likely for /d/ than for /b/ or /g/. Lovins points out that /d/, unlike /b/ or /g/, takes an epenthetic /o/ when it occurs word-finally, probably because underlying coronal stops occurring before /ɯ/ are realized phonetically as affricates [ts] and [(d)z]. Since /o/ is inherently longer in duration than /ɯ/, Lovins suggests that singleton /b/ or /g/ before /ɯ/ is more likely to be judged by a Japanese speaker as being perceptually similar to a word-final /b/ or /g/ in English than a singleton /d/ occurring before /o/, and thus /d/ tends to be geminated more than /b/ or /g/. Katayama, on the other hand, suggests that the difference stems from the inherent duration differences in the stops themselves. Since /d/ tends to be shorter in duration than /b/ or /g/, it will be the easiest of the three consonants to geminate. However, while both accounts can explain why /dd/ is more acceptable than /bb/ or /gg/, neither seems to be able to explain why /gg/ is also more acceptable than /bb/, since they both assume that /bb/ and /gg/ pattern together for the purposes of degemination, when in fact the historical data suggests that /gg/ should pattern with /dd/.

A third possibility is that the relative acceptability of geminates stems from the difficulty of maintaining voicing over a long period of time (Ohala 1983, Hayes & Steriade 2004). This would predict that /dd/ should be more acceptable than /gg/, because the distance from the glottis to the oral closure is longer in /dd/ than in /gg/, making it easier to maintain voicing in /dd/. However, this account would then predict that /bb/

should be even more acceptable than /dd/, since the distance from the glottis to the lips is even longer than the distance from the glottis to the alveolar ridge. Yet historically word-final /d/ or /ʤ/ is the most likely to be geminated in Japanese loanwords, while word-final /b/ is hardly ever geminated at all.

I suspect that the relative unacceptability of /bb/ is not due to a phonetic reason, but rather is because of the unusual nature of /p/ and /b/ in Japanese phonology. Recall in Chapter 1, in my discussion of the markedness constraint *P, that singleton /p/ rarely occurs in native Japanese words (other than mimetics), due to a historical lenition of Old Japanese *p. This has resulted in /b/ patterning with /h/ instead of /p/ in voicing alternations such as *rendaku*. If historical Japanese speakers based their acceptability judgements of voiced geminates on the likelihood of their voiceless geminate counterparts occurring in native Japanese words, then they would have considered /bb/ less acceptable than /dd/ or /gg/, since /hh/ never occurs in Yamato or Sino-Japanese, and /pp/ is rare, while /tt/ and /kk/ are quite common. The differing numbers of loans containing geminate /bb/, /dd/, or /gg/ may also have been a factor in determining the modern-day gemination pattern. Figure 2.6 shows that among the loanword data I collected from Arakawa (1977), words derived from English word-final /d/ are the most common, followed by /g/, then /ʤ/, then finally word-final /b/ is the least common. This suggests that the higher type frequency of loans with final /d/ or /g/ had an effect in making /dd/ and /gg/ more acceptable than /bb/. In Chapter 4 I will further address the role of frequency in determining these adaptation patterns by developing a connectionist model of the adaptation of voiced geminates, showing that the marked nature of /p/ in Japanese phonology is one of the factors contributing to the relative unacceptability of /bb/.

Table 2.21: Adaptation of *team*→[ʧiːmɯ], *[tiːmɯ]

tim	*TI	FAITH-SB
a. tiːmɯ	*!	
b. ☞ ʧiːmɯ		*

2.3.5 Coronals and palatals before front vowels

The final example of loanword adaptation that I will present in this section concerns the distribution of coronal and palatal obstruents before front vowels in Japanese loanwords. As was mentioned in section 2.2, the coronal obstruents [t], [d], [s], and [z] do not occur before /i/ in Yamato, Sino-Japanese, and Mimetic words. This statement holds not only as a static generalization over possible word forms in the (non-Foreign) lexicon, but also governs morphophonemic alternations in verb conjugation patterns. For example, the final /t/ in the root /mat−/ 'wait' surfaces as [t] in a form like /mat+anai/ → [ma<u>t</u>anai] 'wait (neg.)', but as [ʧ] in /mat+i+masɯ/ → [ma<u>ʧ</u>imasɯ] 'wait (polite)'. Ito & Mester (1995) propose a constraint *TI (coronal stops cannot appear before /i/) to account for this pattern.[34] On the other hand, the palatal[35] obstruents [ʧ], [ʤ], and [ʃ] do not occur before /e/. However, this is only a static generalization over the lexicon; there are no morphophonemic alternations in which an underlying /ʧ,ʤ,ʃ/ surfaces as [t,d,s] before /e/. Ito & Mester (1995) propose another constraint *ČE ruling out palatal obstruents before /e/. Some example tableau showing the effects of *TI and *ČE are shown in Tables 2.21 and 2.22.

Because the *TI and *ČE constraints do not hold in English, there are many po-

[34]Ito & Mester (1995) also propose *SI, which disallows coronal fricatives before /i/. The reason they propose two separate constraints instead of a more general one holding for all coronals is the existence of loans which violate *TI but not *SI, for example シティバンク /ʃitibaɴkɯ/ 'Citibank', where only /s/, and not /t/, is palatalized before /i/.

[35]For convenience I refer to these segments as palatals, but in fact they are more properly described phonetically as alveolo-palatals (Akamatsu 1997).

Table 2.22: Adaptation of *shepherd*→[sepaːdo], *[ʃepaːdo]

ʃɛpɹ̩d	*ČE	FAITH-SB
a. ʃepaːdo	*!	
b. ☞ sepaːdo		*

tential loans which would violate these constraints if they were borrowed directly into Japanese. Since both /i/ and /ɪ/ in English are usually adapted as /i/ in Japanese, any word containing /t/, /d/, /s/, or /z/ before either of these two vowels is a potential *TI violation. Likewise, any English word with /ʧ/, /ʤ/, or /ʃ/ occurring before /e/ or /ɛ/ is a potential *ČE violation. Among the potential *TI violations, these loans generally show three different possible adaptations of the illegal coronal-/i/ sequence:

ČI Palatalize the coronal, giving [ʧi, ʤi].

TE Lower /i/, giving [te, de].

TI Make no changes, giving [ti, di] (and violating *TI).

Some representative potential *TI violations from English, and attested forms of the corresponding loans from Arakawa (1977), are presented in Table 2.23. Potential *ČE violations likewise have three possible adaptations:

SE Depalatalize /ʧ,ʤ/, giving [se,ze].

ČI Raise /e/, giving [ʧi,ʤi].

ČE Make no changes, giving [ʧe,ʤe] (and violating *ČE).

The question arises as to what determines whether a given loanword obeys *TI/*ČE or not. With *ČE violations, the adaptation used depends on the source obstruent, with ČE being used consistently for source /ʧ/ in all time periods (except for セロ /seɾo/

Table 2.23: *TI violations and attestations given in Arakawa (1977)

Source word	Spelling	Pronunciation	Date	Adaptation
ticket	チケット	[ʧiketto]	1867	ČI
	チケット	[ʧiketto]	1869	ČI
	チケット	[ʧiketto]	1881	ČI
	ティケット	[tiketto]	(??)	TI
	ティケット	[tiketto]	(??)	TI
society	ソサイティ	[sosaiti]	1871	TI
	ソサイティ	[sosaiti]	1874	TI
	ソサイテー	[sosaiteː]	1884	TE
	ソサイテー	[sosaiteː]	1889	TE
	ソサイエチー	[sosaieʧiː]	1889	ČI
lady	レーディ	[reːdi]	1878	TI
	レディ	[redi]	1880	TI
	レデー	[redeː]	1883	TE
	レディ	[redi]	1885	TI
	レジー	[reʤiː]	1886	ČI
	(レディー)	[rediː]	1887	TI
dilemma	ディレンマ	[direɴma]	1896	TI
	ディレンマ	[direɴma]	1907	TI
	(ジレンマ)	[ʤireɴma]	1910	ČI
	(ジレンマ)	[ʤireɴma]	1911	ČI

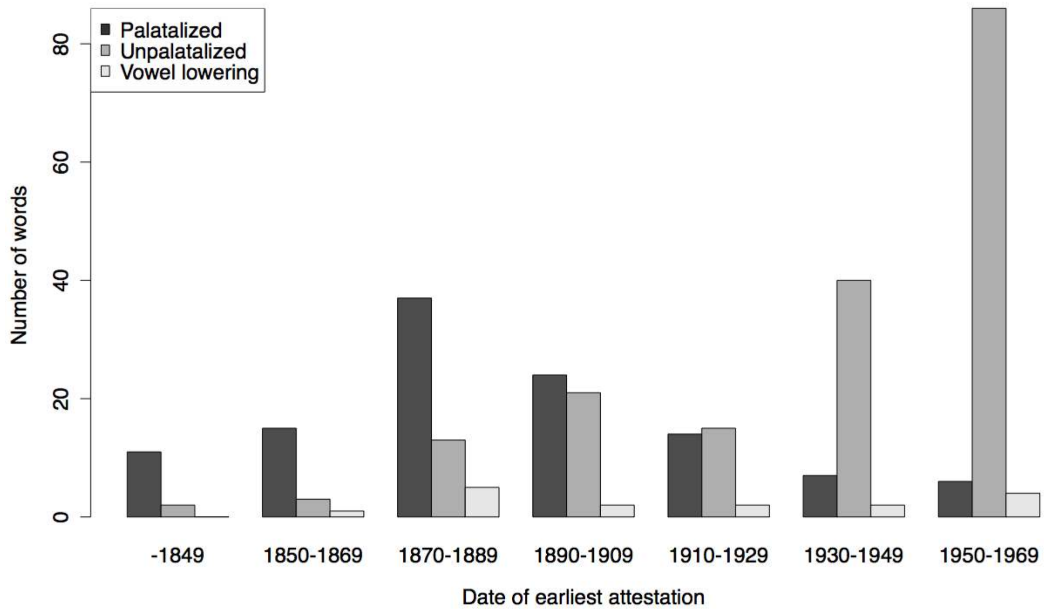


Figure 2.7: Adaptation patterns over time of *TI-violating loanwords

'cello'), and SE or ČE for /ʤ/ and /ʃ/, with both of the latter tending to be adapted as ČE only for words borrowed after about 1950. The ČI adaptation is practically never attested for these loanwords; the only exceptions from Arakawa (1977) are チッキ [ʧikki] for *check*, competing with the more common adaptations チェッキ [ʧekki] and チェック [ʧekkɯ]; and a nonce adaptation from 1871 of ジェズイット [ʤezɯitto] 'Jesuit' as ジスイト [ʤisɯito].

However, the *TI violations show a more complex pattern of variation in adaptation patterns. As with the velar palatalization and voiced geminate adaptation patterns, it turns out that the age of the loan is a factor here. I collected 339 loanwords from Arakawa (1977) derived from source words containing a coronal stop followed by [i] or [ɪ]. These loans were then classified by the most common adaptation strategy used in each word's cited forms. This data is summarized in Figure 2.7. One thing that is immediately striking is that the total number of new loanwords coming into the language has generally increased over time, with two discernible peaks occurring in the periods

1870–1889 and 1950–1969. These two peaks correspond with two significant events in Japanese history which increased Japanese contact with the outside world: the beginning of the modernization of Japan during the Meiji era, and the post-WWII occupation of Japan by the United States, respectively. As far as the individual adaptation patterns themselves, words first attested before about 1890 usually have the ČI adaptation, while words first attested after about 1930 usually have TI instead, with a gradual shift from the TI→ČI to the TI→TI adaptation strategy taking place from 1870–1930. The third adaptation pattern, TI→TE, where the coronal stop is preserved but the vowel is lowered to [e], turns out to never be very common relative to the other two adaptation patterns, except for a slight peak during the period 1870–1889.

While tokens of early *TI-violating loans which generally show the ČI adaptation are sometimes attested with the TI adaptation around the time they are first borrowed, tokens of more recent loans are almost always attested with the TI adaptation only, and these words show no sign of becoming more nativized over time. This suggests that there has been a change in the acceptability of TI sequences among Japanese speakers over time: mostly unacceptable before about 1890, variably acceptable from 1890–1930, and mostly acceptable after about 1930. What's more, this change in acceptability seems to be sensitive to the voicing of the stop in the source word, with loanwords derived from source words containing /t/ being more likely than loans derived from words containing /d/ to be nativized over the time period we are looking at. This can be seen by splitting up the loanword data by the voicing of the stop, as in Figures 2.8 and 2.9. The [ti,tɪ]→[ʧi] adaptation strategy was generally preferred over [ti,tɪ]→[ti] until about 1930, with a period of variation between [ʧi] and [ti] adaptations occurring from 1890 to 1930 (Figure 2.8). The [di,dɪ]→[di] adaptation strategy, on the other hand, came to be preferred over [di,dɪ]→[ʤi] at an earlier time, between 1910–1930, with variation between [ʤi] and [di] from 1890–1910 (Figure 2.9). The [di,dɪ]→[de]

adaptation strategy was also more common than [ti,tɪ]→[te] in the late 19th and early 20th centuries, although it never ultimately became the preferred adaptation for either set of TI sequences.

There also seems to be a kind of phonological neighborhood affect governing which adaptation strategy is likely to apply to a given loanword. For example, while words with /−t+ɪv/ in English (such as *active* or *creativity*) follow the general pattern for voiceless /t/ as being adapted as [ʧi] before 1930 and [ti] afterwards, words with final /−ti/ (such as *city* or *humanity*) or /−t,d+ɪŋ/ (such as *wedding* or *batting*) were commonly adapted as [ti] much earlier, starting around 1870 (Figure 2.10), while words with /−tɪk(s)/ (such as *statistics* or *mystic*, all containing the quasi-morphemic suffix −*ic*) were variably adapted as either [ʧi] or [ti] from 1900–1930 (Figure 2.11). In general, the set of environments in which TI sequences were acceptable in loanwords has gradually expanded over time: at first [ti] and [di] were preserved only if they occurred in the source word in a word-final syllable with /i/ or /ɪŋ/, then [di] from word-initial syllables and [ti] from word-final syllables with /ɪk(s)/ also began to be preserved, and finally [ti] and [di] became acceptable in all other environments, which is the present day situation.

It turns out that the main factor predicting how likely nativization is in a particular phonological neighborhood is the type frequency of the neighborhood among all loanwords with TI. This is shown in Table 2.24, which lists the most common phonological neighborhoods among the TI loans collected from Arakawa (1977) by their date of first attestation. Comparing this table with Figures 2.10 and 2.11, we can see that TI sequences in the most common phonological neighborhoods came to be adapted as TI instead of ČI earlier than the TI sequences in other phonological neighborhoods. Loans with /−ti#/ and /−tɪk(s)/ tend to be the most common in terms of type frequency

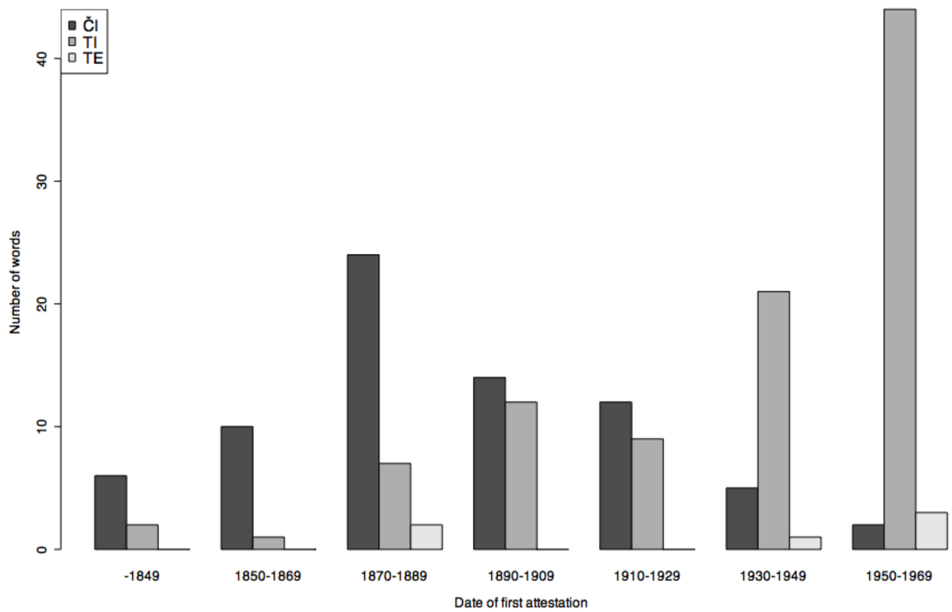


Figure 2.8: Adaptation patterns over time of source words containing /t/

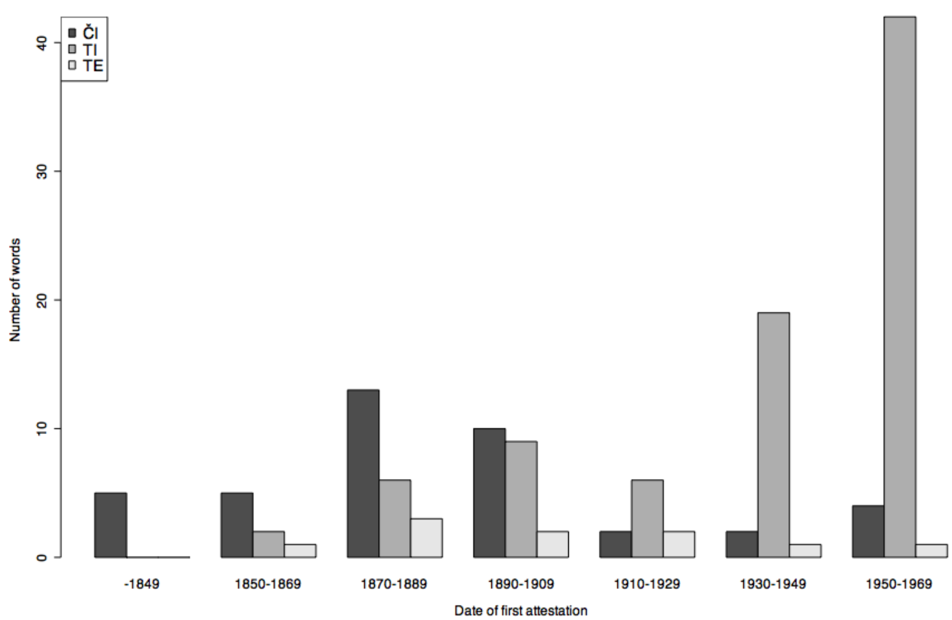


Figure 2.9: Adaptation patterns over time of source words containing /d/

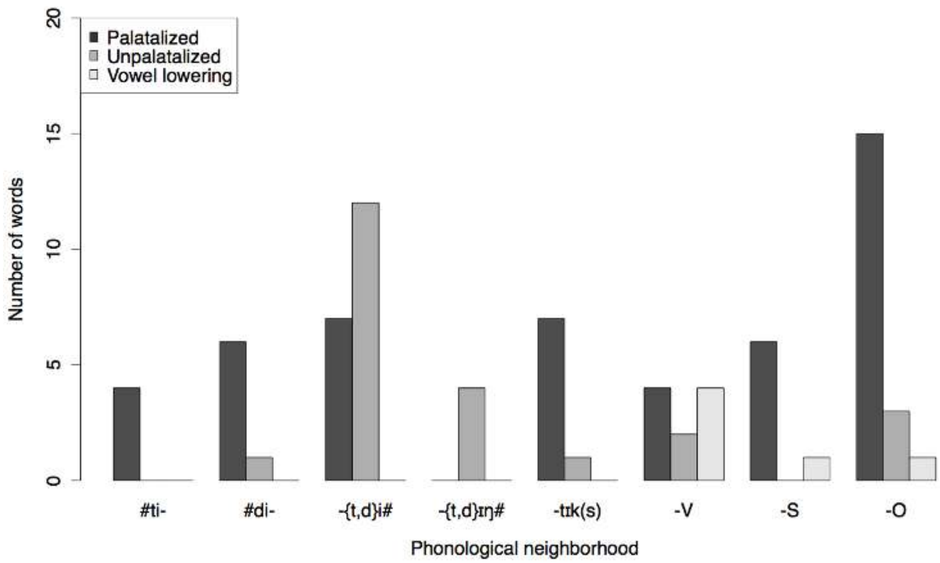


Figure 2.10: Adaptation patterns by phonological neighborhood for TI sequences in loanwords first attested 1870–1899. V=vowel; S=sonorant; O=obstruent.

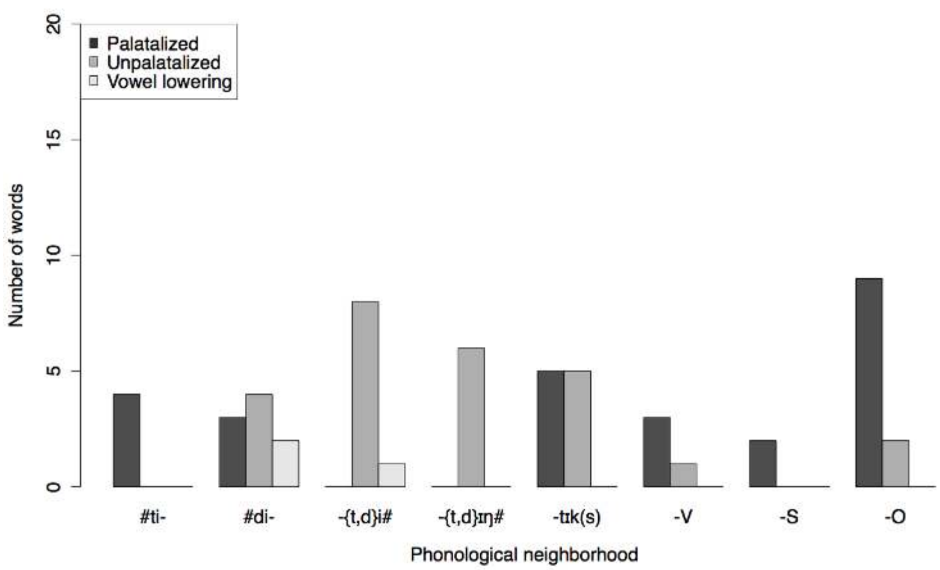


Figure 2.11: Adaptation patterns by phonological neighborhood for TI sequences in loanwords first attested 1900–1929.

Table 2.24: Type frequencies of phonological environments for TI-containing loanwords. (Only environments with 5 or more examples for each time period are listed here.)

1870–1889		1890–1909		1910–1929		1930–1949		1950–1969	
−ti#	14	−ti#	8	−tɪk(s)	8	−tɪk(s)	9	−tɪŋ#	16
−tɪk(s)	6	#d−	8	−ti#	5	#d−	7	#d−	14
#d−	6	−tɪk(s)	6			−diV	6	−diV	11
−diV	5	−diV	6					−ti#	9
−di#	5							−di#	6
								#dɪs−	6
								−dɪŋ#	5
all /ti/	40	all /ti/	28	all /ti/	23	all /ti/	30	all /ti/	50
all /di/	25	all /di/	23	all /di/	12	all /di/	24	all /di/	48

among loans first attested from 1870–1930, and these were also the first environments in which the TI adaptation is attested. Loans with /#d−/ and /−diV/ are the next most common in this time period, and loans with /#d−/ started to be adapted as TI from 1900–1930, before most other TI loans.[36] The only anomalous patterns are word-final /−di#/ and /−{t,d}ɪŋ#/, both of which are not all that frequent and yet are usually adapted as TI starting in 1870–1900; and /−diV/, which is about as common as /#d−/ and yet continues to be commonly adapted as ČI well into the 20th century, as shown by loans like ラジウム /raʤiɯmɯ/ 'radium', dating from 1904, and ラジオ /raʤio/ 'radio', from 1926. There are probably some phonetic factors affecting different phonological neighborhoods as well; for example, the /−di#/ pattern is perceptually very similar to the /−ti#/ pattern, since the stop would generally be flapped in both cases in dialects of American English, while obstruents in word-initial position tend to be more perceptually salient than word-medial ones, which might explain the different behavior of /#d−/ and /−diV/. At any rate, there does seem to be a correlation between the type

[36]There happen to be two interesting nonce adaptations attested in Arakawa (1977), サジスティック /saʤisɯtikkɯ/ 'sadistic', dating from 1953, and オプチミスティック /opɯʧimisɯtikkɯ/ 'optimistic', from 1963, which preserve the /t/ in the /−tɪk/ environment from the source word even though the other TI sequence in the word is nativized. These two loanwords are usually attested as サディスティック /sadisɯtikkɯ/ and オプチミスチック /opɯʧimisɯʧikkɯ/, respectively.

Table 2.25: The adaptation pattern TI→TI, with {IDENT-SB(cont), IDENT-SB(high)} ≫ *TI

ti	IDENT-SB(cont)	IDENT-SB(high)	*TI
a. ☞ ti			*
b. ʧi	*!		
c. te		*!	

Table 2.26: The adaptation pattern TI→ČI, with {*TI, IDENT-SB(high)} ≫ IDENT-SB(cont)

ti	*TI	IDENT-SB(high)	IDENT-SB(cont)
a. ti	*!		
b. ☞ ʧi			*
c. te		*!	

frequency of the phonological environment that a TI sequence occurs in and the likelihood that it will be nativized over time, with TI sequences occurring in more frequent phonological neighborhoods being less likely to be nativized to ČI.

Adaptation of TI

Among the TI loans there are three possible outcomes, as I discussed above: TI, ČI, and TE. The TI→TI adaptation pattern can be obtained by ranking FAITH-SB constraints above *TI, as in Table 2.25. In the TI→ČI pattern, the stop /t/ is palatalized and changed to an affricate, meaning that *TI must be ranked above the faithfulness constraint IDENT-SB(cont) preserving the input values of [cont] features (Table 2.26). Finally, TI→TE comes about by ranking *TI above IDENT-SB(high), as in Table 2.27.

However, the constraint *TI is not specific enough if we want to also account for the changes in TI adaptation patterns over time. This is because the TI loans did not

Table 2.27: The adaptation pattern TI→TE, with {*TI, IDENT-SB(cont)} ≫ IDENT-SB(high)

ti	*TI	IDENT-SB(cont)	IDENT-SB(high)
a. ti	*!		
b. ʧi		*!	
c. ☞ te			*

Table 2.28: Adaptation of *director*→[direkɯtaː], with *TI ≫ IDENT-SB(cont) ≫ *DI

dɹɛktɹ̩	*TI	IDENT-SB(cont)	*DI
a. ʤirekɯtaː		*!	
b. direkɯtaː			*

all switch at once from the TI→ČI to the TI→TI adaptation pattern; instead TI→TI occurred initially in very restricted contexts, then was gradually extended to more and more phonological neighborhoods over time until the present day, where TI is the default adaptation for new loans. To do this under OT, the constraint *TI will need to be split into a set of more specific constraints ruling out TI sequences in various environments. For example, suppose *TI is split into two more specific constraints sensitive to the voicing of the stop, *TI (no /t/ before /i/) and *DI (no /d/ before /i/), which can then be ranked independently of each other with respect to IDENT-SB(cont). Then the relative acceptability of [di] (as compared to [ti]) in loans borrowed during the early 20th century can be derived using the ranking *TI ≫ IDENT-SB(cont) ≫ *DI, which would allow [di] to be adapted as [di], but force [ti] to be adapted as [ʧi] (Table 2.28).

The distinction between these two markedness constraints, *TI and *DI, is not motivated by any native phonological processes, however, and would exist only to explain the loanword data. This is similar to Ito & Mester's (1999) postulation of the two constraints *TI (no coronal stops before /i/) and *SI (no coronal fricatives before /i/), instead of a single constraint banning all coronal segments before /i/, which would adequately account for the native distribution of coronals. This would be an unwelcome complication of the theory developed so far. Up until this point, the only loanword-specific constraints needed were the SB-Correspondence constraints, and it was not necessary to postulate loanword-specific markedness constraints as well (except possibly *BB for voiced geminates). As well, to account for the more fine- grained phonological neighborhood effects in adaptations, there would need to be even more specific constraints, such as *#TI (no word-initial /t/ before /i/) or *DIŋ (no /d/ before the sequence /ɪŋ/), which are not only even less plausible than *TI and *DI, but also refer to non-native features of the input, such as whether the vowel is tense or lax, or whether the TI sequence occurs with the segment /ŋ/ (which is not a contrastive phoneme of Japanese). Besides this, there is no way to directly represent the effect of type frequency on TI adaptation patterns using OT constraints. I will return to this issue in Chapters 3 and 4, arguing that the above provides evidence for a model of lexical representation in which the frequency and phonological similarity of lexical entries can affect their processing.

2.4 General observations

Although the attestation patterns for the three groups of loanwords discussed in the previous section are quite complex, in each case the set of attested adaptation patterns can be generated by reranking IDENT-SB and ALIGN-SB constraints against a fixed ranking of markedness constraints, and with MAX-SB $\gg$ DEP-SB (ensuring that epenthesis is favored over deletion for repairing illegal clusters and coda consonants). In the TI case, the three adaptations, TI, ČI, and TE, can be generated with various rankings of IDENT-SB(cont) and IDENT-SB(high) with respect to *TI. The KYA, KA, and KE adaptations for KÆ loans involve reranking IDENT-SB(back) and IDENT-SB(low) with respect to *C^ja, although here the difference between KYA and KA/KE can also be obtained with differences in the palatalization of the input velar stop. The various adaptations for voiced geminates, meanwhile, can be generated with different rankings of IDENT-SB(voi) and ALIGN-SB with *DD. Thus an OT analysis using SB-Correspondence constraints can be used to generate a typology of possible adaptations available to Japanese speakers at the beginning stages of contact with English.

Yet it is more difficult using OT constraints to also account for the cognate effect in KÆ loans, as well as the type frequency and phonological similarity effects on the rate of use in voiced geminates and TI loans of the possible adaptation patterns generated by the grammar. Concerning this latter issue, there are some general trends that can be discerned. Figures 2.12 and 2.13 summarize the nativization data from two of the three groups: coronals and palatals before front vowels, and voiced geminates. (Palatalization of velars in loanwords was not included because unlike the other two examples, it is not an example of nativization; both plain and palatalized velars can occur before /a/ in native words as well as loanwords.) In these figures, coronals before /i/ and palatals before /e/ are graphed separately, and I have also included a fourth set of loanwords, those containing /ɸ/ occurring before a vowel other than /ɯ/. (Recall from Chapter 1 that in native words, [ɸ] is an allophone of /h/ that occurs only before /ɯ/.) In Figure 2.12, for each loanword, the ratio of all non-native tokens to total tokens attested in Arakawa (1977) is plotted by the word's date of first attestation. For example, in the entry for the word /tiːsɯpɯːɴ/ 'teaspoon', Arakawa gives one example where it is spelled ⟨ʧi⟩, and three where it is spelled ⟨ti⟩, giving an F/Total ratio of 0.25. In Figure 2.13, however, only the tokens attested within each 20-year timespan are counted. In the *teaspoon* example above, the ⟨ʧi⟩ token is the earliest attested one, dating from 1867, but the other three date from after 1935, so in this case only the ⟨ʧi⟩ token would count for Figure 2.13, giving an F/Total ratio of 0 for this word. Although this makes the data

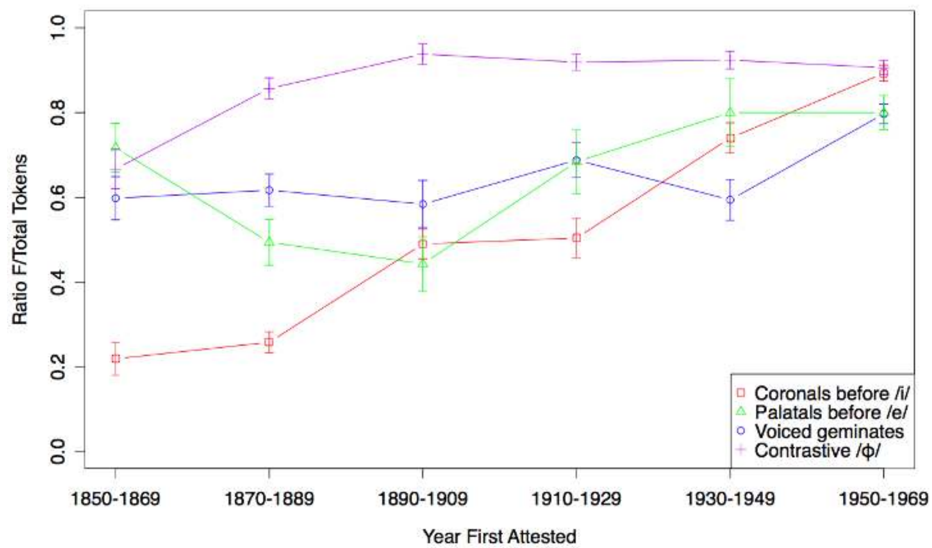


Figure 2.12: Estimated nativization rates (as measured by the ratio of non-native tokens to total tokens attested) by date of first attestation in Arakawa (1977) for coronals before /i/, palatals before /e/, voiced geminates, and contrastive /ɸ/. Error bars indicate one standard error.

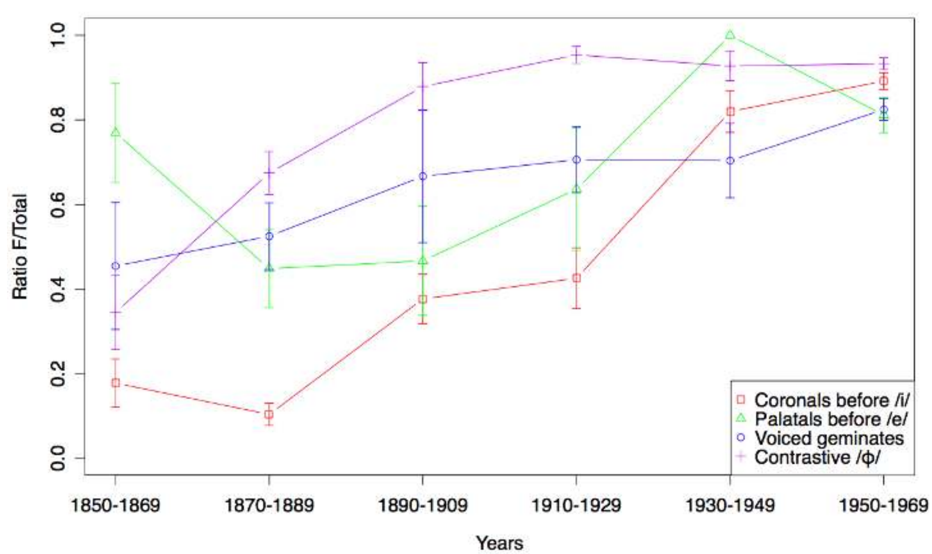


Figure 2.13: Estimated nativization rates (as measured by ratio of non-native tokens to total tokens for 20-year period in which loan is first attested) for coronals before /i/, palatals before /e/, voiced geminates, and contrastive /ɸ/. Error bars indicate one standard error.

somewhat more noisy, since there are fewer overall tokens being used, computing the nativization rates this way makes it easier to relate this data to the model of transmission developed in Chapter 5. In both figures, the error bars indicate one standard error, as calculated for proportions ($S.E. = \sqrt{\frac{p(1-p)}{n}}$, where p is the F/Total ratio for each loanword group, and n is the total number of tokens used in calculating each ratio).

In both figures, there is a general trend towards less nativization over time for all loanword groups. This is most prominent for coronals before /i/, which increases from an F/Total ratio of about 0.2 in 1850–1869 to 0.8 in 1950–1969. This rise seems to begin in the period 1870–1889, when contact with the West was reestablished, and numerous numbers of loanwords entered into the language. After 1890, the nativization rate for TI is comparable to that of other loanword patterns.

There is also an effect of type frequency, which was discussed in the sections on voiced geminates and coronals before /i/. In general, non-native phonotactic patterns are attested earlier in phonological neighborhoods with a high type frequency, then gradually spread to less frequent neighborhoods over time. This gradual generalization of adaptation patterns over time is reminiscent of lexical diffusion in sound change (Wang 1977, Labov 1981).

2.5 Conclusion

In this chapter I have analyzed the borrowing of a new loanword by an L1 speech community as consisting of two separate processes: the adaptation of the word from an L2 source by one or more L1 speakers, followed by the transmission of the loanword to other L1 speakers. I then argued that it is not only the initial borrowers, but also the L1 speakers who are involved in transmission, that are able to perform nativizations and influence the eventual established form of the loanword. This is supported by both theoretical arguments about the cumulative effects of errors in production and perception during the process of transmission, as well as empirical evidence from Poplack et al. (1988) showing that the degree of nativization seen in a loanword correlates with the number of speakers who are attested using it. I then examined three adaptation patterns in detail, the palatalization of velars occurring before a source word /æ/, the gemination of word-final voiced obstruents, and the distribution of palatals and coronals before front vowels, focusing on what historical attestation patterns can tell us about the interplay of adaptation and transmission in the establishment of these adaptation patterns over time. The main finding is that the set of possible adaptations for each pattern can be generated by reranking FAITH-SB constraints with respect to the fixed ranking of markedness constraints used in the non-loan phonology.

In the next three chapters I will examine the processes of adaptation and transmission in more detail. In Chapter 3, I will look at various models of adaptation and lexical representation of loanwords that have been proposed in the generative literature, focusing in particular on the Core-Periphery model of Ito & Mester (1995, 1999). Then, in Chapter 4, I develop a connectionist implementation of a model of the lexicon which can account for the phonological neighborhood and type frequency effects seen in adaptation patterns. In Chapter 5 I then develop a formalized framework for characterizing the effects on the process of nativization of loanword transmission among a network of speakers.

CHAPTER 3

GENERATIVE MODELS OF LOANWORD ADAPTATION AND REPRESENTATION

In the last chapter I argued that loanword borrowing, which is best viewed as a process involving an entire speech community (and not just a single speaker), can be broken down into two distinct processes, namely the *adaptation* of a new word by L1 speakers in contact with L2, followed by the *transmission* of the new loanword from these initial borrowers to the rest of the speech community. In this chapter I focus on the adaptation stage of loanword borrowing. Based on Silverman (1992) and Yip (1993, 2006), this stage of the borrowing process can be further broken down into two sub-processes: the *segmental parse* (SP) stage, taking place during the perception of a new loanword, and the *constraint satisfaction* (CS) stage, taking place during the production of a loanword token (Figure 3.1). The SP stage involves the mapping of an L2 source word by an L1 listener to a string of L1 segments, and is constrained by the perceptual biases of the L1 listener, as well as the set of segments and phonemic contrasts available in the inventory of L1. The L2 source may be either a surface phonetic string as produced by an L2 speaker, an L2 orthographic representation of the word, or, possibly, an L2 phonological representation (although this last possibility is less likely, as I will discuss).

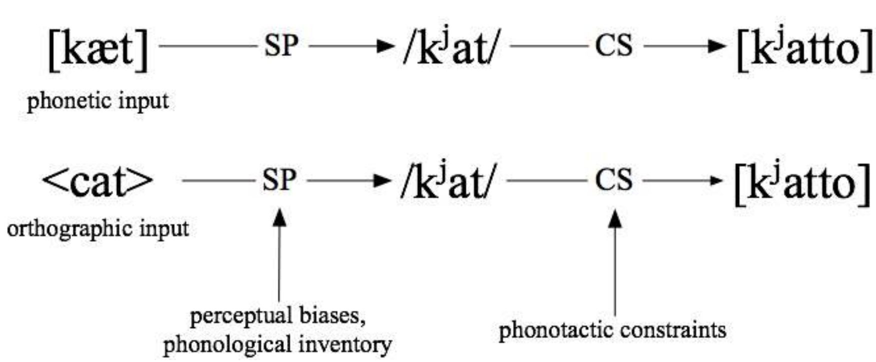


Figure 3.1: The SP and CS stages in loanword adaptation

The output of the SP stage is a string of L1 segments that may possibly violate L1 phonotactics. This string is then submitted to the CS stage, where it is evaluated against a set of L1 phonotactic constraints to produce a target for production.

The main issues I focus on in this chapter are the following:

1. Is the input to adaptation phonetic, phonological, or orthographic in nature?
2. When are *nativizations* (changes in the loanword to make it conform more closely with L1 phonology) made: during the SP stage, the CS stage, or both stages? How does the nature of the input modality (phonetic, orthographic, or phonological) in which an L2 word is perceived affect nativization?
3. How are loanwords represented in an L1 speaker's lexicon? What do loanword adaptation patterns tell us about the structure of L1 speakers' lexicons?

I begin by discussing various proposals regarding issues 1 and 2 above, arguing that many of these make assumptions about the nature of the input and the stage at which nativization takes place which do not hold for all possible borrowing situations. I then discuss in detail the *Core-Periphery* model of the lexicon proposed in Ito & Mester (1995, 1999), focusing on the issue of whether loanwords form a distinct stratum in the Japanese lexicon.

3.1 The input to loanword adaptation

A major debate in the literature on loanword borrowing has been over whether adaptations are based directly on the L2 phonemic representation of the source word (which is then further adjusted to conform to L1 phonology), or is instead based on the surface phonetic form of the source word as it is produced by an L2 speaker (which is then parsed according to the perceptual mechanisms of native L1 listeners). A third option, one which is rarely acknowledged, is that the input is an L2 orthographic representation which is then "sounded out" by an L1 speaker into a string of L1 phonemes.

The first two viewpoints above entail different assumptions about who is performing the initial adaptation of a new loanword. The phonology-only viewpoint assumes that an adaptation is performed by L1/L2 bilingual speakers during the production of a new loanword, by directly mapping the L2 phonological representation of the word to an L1 segmental representation. Nativizations then would be the result of changing the L1 representation to obey L1 phonotactics. Recent proponents of this view include Ito & Mester (1999), Paradis & LaCharité (1997), Jacobs & Gussenhoven (2000), and LaCharité & Paradis (2005), and according to Hyman (1970), is similar to the views of many of the American structuralists as well, for example Bloomfield (1933), Haugen (1950), and Weinreich (1953, 1957). The perception-only approach instead assumes that a loanword adaptation is performed by an L1 listener perceiving a phonetic token of the L2 word being produced by an L2 speaker, with nativizations being the result of misperceptions due to the differences in L1 and L2 phonetics and phonology. This view is argued for in recent work such as Kang (2003), Iverson, Kuhl, Akahane-Yamada, Diesch, Tohkura, Ketterman & Siebert (2003), Peperkamp & Dupoux (2003), Peperkamp (2004), and Iverson & Lee (2006), among others, although Hyman (1970) shows that it dates back at least to Neogrammarians such as Paul (1880/1889).

Evidence can be found supporting both points of view. On the perception-only side are cases where an allophonic contrast in L2 is apparently perceived as an L1 phonemic contrast and is preserved as such in loanwords. For example, Cantonese makes a contrast between aspirated and unaspirated stops, and Cantonese speakers consistently adapt word-initial voiceless stops from English as aspirated (e.g. *tie*→[t^haj]), but voiceless stops in /sC/ clusters as unaspirated (*stick*→[sitik]; Silverman 1992). This reflects the allophonic difference in aspiration for voiceless stops in word-initial and /sC/ cluster contexts, and thus shows that Cantonese speakers are basing their borrowings on the surface phonetic forms of the English source words, not on the underlying phonological representations. Hawaiian Japanese provides another example of a perceptually-motivated nativization. English words with coronal stops which occur in a flapping environment (intervocalically, after a stressed syllable) are sometimes borrowed as alveolar flaps into this dialect of Japanese, for example in *thirty*→[toːɾi] (Higa 1970).[1] Since *thirty* does not have /ɾ/ underlyingly (/ɾ/ is not even a member of the phonemic inventory of English, after all), the input for this borrowing had to have been the surface form, not the phonological representation.

Yet there are also examples of adaptation patterns which were clearly *not* based on the surface L2 phonetic form. While these examples are usually considered to be phonologically-based adaptations, in fact it is hard to tell in many of these cases whether the input was L2 phonology or orthography. An example comes from Korean, which, like Cantonese, makes a distinction between aspirated and unaspirated stops. Yet voiceless stops from English are always adapted into Korean as aspirated, even in /sC/ clusters: *tie*→[t^{h}ai], *stick*→[sɨt^{h}ik]. Oh (1996) suggests that examples like these show that Korean borrowers are using as the inputs for adaptation the phonological representation of these words, in which *tie* and *stick* both have /t/ underlyingly, and there is an adaptation rule used by Korean speakers which maps English /t/ in all cases to Korean [t^{h}]. But, since /t/ is spelled with a ⟨t⟩ in both *tie* and *stick*, in this case it is just as possible that Korean borrowers are using an orthographically-based adaptation strategy, such as

[1] Note that coda /r/s are deleted in loanwords in this dialect, as in Standard Japanese (for example, in *Thursday*→[toːzɯde]; the Standard Japanese equivalent, if it exists, would probably be something like *[saːzɯdeː]), so the /ɾ/ in [toːɾi] was derived from the surface [ɾ] in [θɪɹɾi], not from the [ɹ]. However, flapping is inconsistent in Hawaiian Japanese; another example from Higa (1970), *daddy*→[dedi], shows the surface [ɾ] being adapted as /d/ instead.

"always pronounce the letter ⟨t⟩ as [tʰ]", instead of a phonologically-based one.

Another example of a non-phonetic adaptation pattern is Standard Japanese, which, unlike Hawaiian Japanese, does not borrow coronals in a flapping environment as /ɾ/, so that words like *butter* are adapted as [bataː], not *[baɾaː]. Borrowings in Standard Japanese even preserve underlying voicing distinctions which can be lost after flapping, such that words like *writer* and *rider*, which are nearly homophonous for many speakers of American English (both being pronounced as something like [ɹɑɪɾɹ̩]),[2] are still distinguished in Japanese borrowings (ライター /ɾaitaː/ and ライダー /ɾaidaː/, respectively). This is in spite of the fact that /ɾ/ exists in the phonemic inventory of Japanese, and thus something like *[ɾaiɾaː] would be a closer approximation to the surface form of either word as it is normally produced by an American English speaker.[3] Again, this is a case where it is difficult to distinguish between phonologically and orthographically based adaptations on the evidence of the borrowing outcome alone, since the underlying phonological contrasts (the voicing of the stop in a flapping environment) is also represented in the orthography (*wri<u>t</u>er*, *ri<u>d</u>er*).

In fact, it is generally agreed that in the Japanese case, adaptations in recent loans are based on the spelling of the English source word, not on its phonological structure (Lovins 1975). The Japanese contact situation with English is an example of what Thomason & Kaufman (1988) call category (2) borrowing. In this type of borrowing,

---

[2]Note that for dialects in which Canadian Raising occurs, these two forms may differ in the vowel quality of the diphthong. In my own Syracuse, NY dialect, for example, *rider* is pronounced [ɹɑɪɾɹ̩], but *writer* is closer to [ɹəɪɾɹ̩], with a raised diphthong [əɪ]. Even so, both forms still have a surface [ɾ] corresponding to an underlying /t/ or /d/, which would be better approximated phonetically by /ɾ/ in Japanese.

[3]One could argue that these two words may have been borrowed from a non-flapping dialect of English, and this is why the /t/—/d/ distinction is preserved in these two particular words. However, the only standard Japanese borrowings that I know of in Arakawa (1977) in which /t/ or /d/ in a flapping environment has been borrowed as /ɾ/ are プリン /puɾiɴ/ 'pudding' and ジルバ /ʤiɾɯba/ 'jitterbug'. This is a surprising fact, from a perception-only view: given the degree of contact between Japanese and American English after WWII, one would expect many more examples of /t,d/→[ɾ] if Japanese borrowers *always* base borrowings on the surface phonetic form of the source words.

the degree of contact is slightly more intense than that of a superordinate language with an outnumbered subordinate population, where words for a few culturally-specific items may be borrowed, such as *spaghetti* or *kosher* into American English, but borrowing of function words, phonology, or grammatical constructions does not occur (p. 77). However, the borrowing is not as intense as situations where there is a fair degree of bilingualism in L1 among L2 speakers, and where phonology and syntactic features may be borrowed, and even spread to native words, depending on the degree of contact. Several of the examples of category (2) borrowing that Thomason & Kaufman give are situations in which an L1 community is literate in—but do not necessarily speak themselves—a prestigious literary language L2; for example, the influence of Classical Arabic on the modern-day Arabic languages, Urdu, and the Turkic languages spoken by Muslims, and the influence of Sanskrit on many Dravidian languages. The influence of English on Japanese over the past 150 years is comparable to these other examples of language contact between a spoken and a literary language. The majority of Japanese speakers have only a passing familiarity with English from having studied it in school (which is virtually mandatory throughout the country), and English education in Japan focuses on reading and writing skills, almost to the exclusion of speaking and comprehension. Most loanwords, then, especially after about 1890 when mandatory English instruction was first instituted (Loveday 1996), would have come into Japanese via the written route; that is to say, a borrower of one of these words originally saw it in print, then derived a (non-native) segmental parse for the loanword from the orthography of the source word. This can be seen, for example, in spelling pronunciations such as the adaptation of English /ə/ in Japanese loanwords, which is adapted as [a] when the source schwa is spelled ⟨a⟩ or ⟨u⟩ (as in *Christma̲s*→[kɯrisɯma̲sɯ] or *circu̲s*→[saːka̲sɯ]) but as [i], [e], or [o] when spelled ⟨i⟩, ⟨e⟩, or ⟨o⟩, respectively (*tenni̲s*→[teni̲sɯ], *busine̲ss*→[bidʒine̲sɯ], *millio̲n*→[mirio̲ɴ]).

Even if the role of orthography is acknowledged in influencing adaptation patterns, it is clear that if all loanwords in all possible contact situations are assumed to be derived from only a single type of input, either phonetic, phonological, or orthographic in nature, then it will not be possible to give a complete account of all of the cases discussed above. To some extent, this controversy stems from viewing loanword borrowing as a unitary process performed by a single speaker, requiring researchers to pick a side over whether a loanword adaptation is performed by L1/L2 bilinguals with access to L2 phonological representations, or by L1 speakers who misperceive L2 surface forms in terms of L1 phonology. If instead the wider view proposed in Chapter 2 is taken, and loanword borrowing is considered to be a process involving the entire speech community, then it no longer becomes necessary to choose between the perception-only and phonology-only views. Even if the initial borrower(s) of a word are L1/L2 bilinguals and use phonologically-based adaptations (or, more likely, L1 speakers who have some literacy in L2 and use orthographically-based adaptations), if some of the non-native features of the word are left unnativized by these borrowers in their production of the word, then it is still possible for further perceptually-based nativizations to be made by the other L1 speakers involved in the transmission of the loanword. In addition, since there is often more than one speaker at a time adapting a new loanword, these different speakers may prefer either the phonological or phonetic strategies for adaptation, depending on their familiarity with L2, as illustrated in Figure 3.2.

This would suggest that it, if the same source word is borrowed by both L1/L2 bilinguals and by L1 monolinguals, then it should be possible for two different forms of the loanword to be attested, one with phonological/orthographic and the other with perceptual adaptations. In fact, there are several cases of loanword doublets in Japanese, mainly dating from the 19th century, in which one member of the pair shows deletion of coda consonants from the source word, while the other member preserves the source

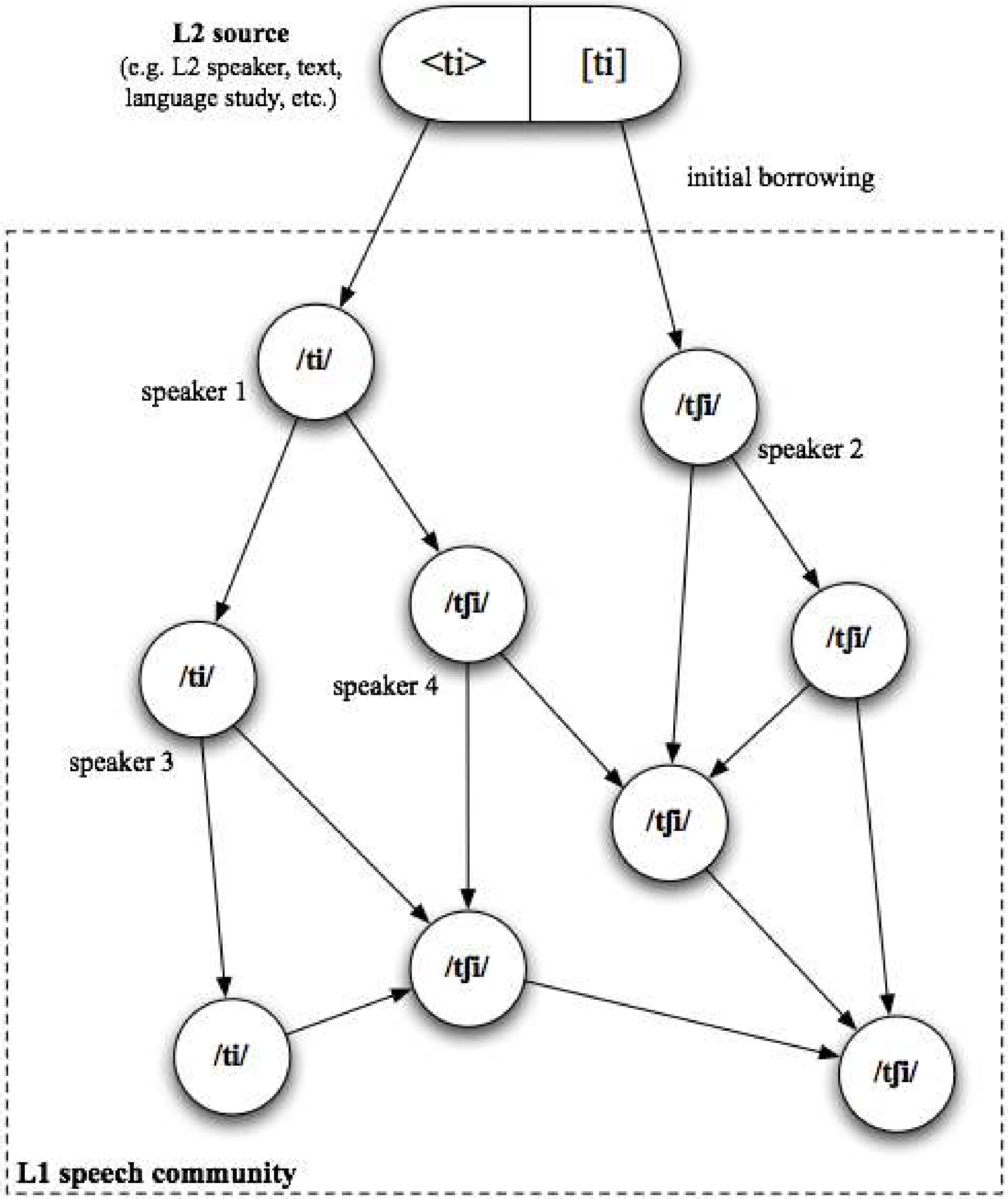


Figure 3.2: Sources of variation in loanword adaptation. Speaker 1 uses a phonological/ orthographic input, adapting the loanword as [ti], while speaker 2 uses a phonetic input, adapting the loanword as [ʧi]. Speakers 3 and 4 use the phonetic output from speaker 1 when learning the loanword, but due to variation in perception end up learning the word as [ti] and [ʧi], respectively.

Table 3.1: Deletion and epenthesis in Japanese loanword doublets, from Smith (2006)

Source word	Deletion form	Epenthesis form
glycerine	ɾi.sɯ.ɾiɴ	gɯ.ɾi.se.ɾiɴ
jitterbug	ʤi.ɾɯ.ba	ʤit.taː.bag.gɯ
lemonade	ɾa.mɯ.ne	ɾe.mo.neː.do
handkerchief	haɴ.ke.ʧi	haɴ.ka.ʧiː.ɸɯ
all right	oː.ɾai	oː.ɾɯ.ɾai.to
don't mind	doɴ.mai	doɴ.to.maiɴ.do

codas via epenthesis (Table 3.1). Smith (2006) argues that the deletion members of each pair are perceptually-motivated, while the epenthesis forms are instead spelling pronunciations, given that the deletion forms often preserve the syllable count of the source words, while in the epenthesis forms, the adaptation of /ə/ depends on the spelling of the source word (the /ə/ in *glyc<u>e</u>rine* is borrowed as /e/, while /ə/ in *lem<u>o</u>nade* is borrowed as /o/).

Thus it is possible for the same source word to be borrowed via both the written and spoken route, showing that either input is potentially available for borrowers, at least in Japanese. These cases of loanword doublets are not that common in modern-day Japanese, however.[4] There seems to be an overall tendency for a speech community to favor over time one or the other of the two types of input, orthographic or phonetic, even if both are available in the initial stages of contact. Borrowings in Cantonese and Hawaiian Japanese tend to be phonetically based, while borrowings in Korean and Standard Japanese tend to be orthographically based. This raises the question of whether it is possible for purely phonologically-based adaptations to be made, with no influence from L2 orthography. All of the cases in the literature that I know of that are claimed to be examples of phonological adaptation involve speech communities with some degree

[4]Pairs that do survive, such as ラムネ /ɾamɯne/ and レモネード /ɾemoneːdo/, both derived from the word *lemonade*, do so when the two words develop a difference in meaning. In this case, /ɾamɯne/ has come to refer to a carbonated lemon-lime soft drink, while the later reborrowing /ɾemoneːdo/ retains the original meaning 'lemonade'.

of literacy in L2, making it difficult to determine whether the initial borrowers based their adaptations on L2 phonology or L2 orthography (or both). A purely phonological borrowing situation would have to involve borrowers who are bilingual in both L1 and L2, allowing them access to L2 phonological representations, but who are not also literate in L2, ruling out any influence of L2 orthography. A contact situation like this would be extremely unlikely in the Western world, at least in modern times, but perhaps historical examples of this kind of language contact could be found among languages spoken in an area with high linguistic diversity and no native writing tradition before European contact (for instance, in the Pacific Northwest of the United States and Canada). However, such contact situations often result in the formation of pidgins like Chinook Jargon (Silverstein 1972) rather than lexical borrowing as such. Examples of purely phonological borrowing, in which orthographic influence can definitely be ruled out, are difficult to find in the language contact literature.

3.2 Nativization during the SP and CS stages

An issue related to the nature of the input for borrowing is whether nativization takes place during either the segmental parse (SP) stage of loanword adaptation, or instead during the constraint satisfaction (CS) stage. If the input is mapped directly from the L2 phonological representation, as LaCharité & Paradis (2005) and others have argued, then of course perception cannot play a role in affecting the outcome of adaptation, and any nativization that occurs must be during the CS stage only. Likewise, orthographically-based adaptations would not be expected to show the effects of phonetic perception, although it is possible for orthographic inputs to have spelling pronunciations or show unusual adaptations on analogy with similarly-spelled words. If perception is involved, however, then either stage (or even both stages) may be involved, as in the models proposed by Silverman (1992) and Yip (1993). The way to tell which stage is involved in a particular nativization is to see if it corresponds to processes which also occur in the native phonology. Yip (2006) gives an example of an SP-level adaptation from Cantonese, which has /l/ but not /r/ in its native inventory. While word-initial /r/ in English words is adapted as [l], as in *rum*→[lɐm], in clusters /r/ is always deleted, as in *friend*→[fɛn]. Yet /l/ is retained in /Cl/ clusters, but not in /sCl/ clusters: *plum*→[powlɐm], but *spleen*→[sipin].[5] Yip argues that the differential deletion of /r/ and /l/ is the result of the interaction of two factors: whether or not the segment exists in the Cantonese inventory (with [r] being more difficult to perceive than [l], since it is not a native segment), and the environment that the segment occurs in (with the different environments being ranked in a perceptibility scale: $\#\_V \succ \#C\_V \succ \#sC\_V$). Neither of these factors could involve native phonological constraints, since neither the segment /r/ nor consonant clusters exist in Cantonese phonology. Yip concludes that these adaptations must occur during the SP stage.

OT accounts of loanword phonology typically do not distinguish between SP-level and CS-level nativizations. Instead, these accounts make one of two claims: that all nativizations take place during the SP stage only (the perception-only view from the previous section), or that they take place during the CS stage only (the phonology-only view). Yet nativizations actually can take place during either stage (as we saw in section 3.1), as some nativization patterns are the result of L1 speakers not being able to perceive a contrast made in L2, or conversely reinterpreting an allophonic alternation in L2 as a phonemic one in L1, while others are the result of constraints against phonotactic sequences that can occur in L2 but not in L1. The relative influence of phonetics or orthography/phonology then depends on whether the L2 word is borrowed via the oral or the written route (with phonetic information not being present in written borrowings),

[5]The final example is an elicited on-line adaptation from Leci & Poon (2004); there are apparently no attested established loans in Cantonese from /sCl/ source words.

and whether the borrower is bilingual in L2 (with monolingual L1 speakers having only indirect access to L2 phonological representations, via the L2 orthography).

A related issue is the different types of phonological regularities that the constraints in CS are used to account for. In OT, the constraint set typically includes both static generalizations about the existence of segments and phonotactic sequences in the lexicon, and generalizations relating to phonological alternations. Given Richness of the Base (Prince & Smolensky 1993: 209), these two types of generalizations cannot easily be distinguished. For example, Rice (1997) and Ota (2004) point out that many of the constraints that Ito & Mester (1995) postulate, such as *NT (no voiceless obstruents after nasals), are used to account for both morphophonemic alternations such as post-nasal voicing of the past tense marker /−ta/, and static distributions such as the lack of post-nasal voiced obstruents in Yamato roots. This ambiguity between these two types of constraints makes it difficult to say whether borrowers perform nativizations solely on the basis of the static distributional constraints in their native lexicon, or on the basis of morphophonemic alternations, or both.

The data from Chapter 2 show that non-native phonotactic sequences which only violate static distributional constraints are preserved in loanwords earlier than those sequences which are also avoided in morphophonological alternations. Among the static distributions discussed in Chapter 1, *ČE (no palatal obstruents before /e/) has had some influence on adaptation patterns, although only /ʃe/ and /ʤe/ sequences have historically been affected; /ʧe/ is nearly always preserved even in very early loanwords like チェッキ /ʧekki/ 'check', which dates from the 1860's (Arakawa 1977).[6] Meanwhile, *NT and *P have been completely inactive in loanword adaptations. There are

[6]There are some static constraints from Ito & Mester (1995), such as *YE ([j] cannot appear before /e/) and □FU ([h] cannot appear before /ɯ/) which are always respected in borrowings, for example *yellow*→[ieroː], *[jeroː] and *hook*→[ɸɯkkɯ], *[hɯkkɯ]. I don't know why *ČE and *F are violable in loanwords while *YE and □FU are not.

no cases that I know of where a post-nasal voiceless stop is adapted as voiced, or a singleton /p/ is realized as [h] or [b] instead in a loanword. Another static distribution, *F ([ɸ] can only occur before /ɯ/) has sporadically affected the adaptation of /f/ from English and other languages, resulting in borrowings such as *siphon*→[saihoɴ] or *fork*→[hoːkɯ]. Yet these forms almost always coexist with /ɸ/ adaptations ([saiɸoɴ], [ɸoːkɯ]), and after about 1890, nearly all words with /f/ were consistently adapted with [ɸ]. The phonemic alternation involving the palatalization of coronals in verb conjugations, on the other hand, seems to have had a much stronger affect on the adaptation of coronals before /i/, since the adaptation of /ti/→[ʧi] and /di/→[ʤi] persisted well into the early 20th century in some environments. This suggests that for speakers in the late 19th century, the *F and *ČE constraints were not as active as *TI for the purposes of loanword nativization. In other words, non-native sequences like /ʧe/ or /ɸa/ are accidental gaps in the lexicon, and were not necessarily impossible for at least some 19th-century Japanese speakers to produce, while sequences like /ti/ were actively avoided in production, making them much more likely to be nativized. I will return to this matter in Chapter 5, where it will be seen that given the attested numbers of nativized and unnativized loanword tokens, as well as the expected effect of transmission on nativization, 19th-century Japanese speakers were apparently much more likely to nativize TI sequences in loanwords than they were ČE sequences, voiced geminates, or violations of *F.

3.3 Lexical stratification

Given that the output of the SP stage may be a non-native phonological representation, the question arises as to how this output is stored and processed in the lexicon of an L1 speaker. The traditional view has been to mark these words as being somehow special and not subject to native phonological processes, for example with a feature [+foreign] (Saciuk 1969). In a long-term borrowing situation with large numbers of L2 words being borrowed into L1, this can result in the L1 lexicon being organized into several distinct *lexical strata*, one consisting of the native vocabulary, and the others consisting of loanwords from the various other languages that L1 speakers have been in contact with. These lexical strata form natural domains over which phonological processes may be restricted, as shown in Chapter 1. I will now examine two recent proposals for how to represent lexical stratification in OT, arguing that neither can fully account for the gradient nature of historically attested loanword adaptation patterns in Japanese. I will then motivate a model for the Japanese lexicon based on connectionist models of lexical processing, and argue that this model can better explain the phonological neighborhood and frequency effects seen in the changes over time on the adaptation patterns discussed in Chapter 2.

3.3.1 The Core-Periphery model of lexical stratification

Ito & Mester (1999) have proposed the most explicit account of lexical stratification in an OT framework, namely the *Core-Periphery* model of the lexicon.[7] The claim is that the different sets of morphophonemic properties shared by the different strata in Japanese are organized in a "core-periphery" structure, a concept reminiscent of the Prototype Theory of semantic categorization (Rosch 1973, Rosch & Mervis 1975), although Ito & Mester do not explicitly make this connection. The basic idea is that, for any class of objects, for example the class of birds, there are some members, like sparrows, which are core members of the class in that they share prototypical features, such as having wings and feathers and being able to fly. Yet there are more periph-

[7] This model is also referred to as the *Indexed Faithfulness* model in later research (e.g. Inkelas & Zoll 2007), because it postulates a set of faithfulness constraints indexed by lexical stratum.

Table 3.2: Pattern of violations for *P and *NT

	Yamato	Sino-Japanese	Foreign
CODACOND	√	√	√
*P	√	√	violable
*NT	√	violable	violable

eral members, such as penguins and ostriches, which we recognize as being birds even though they do not share all of these properties. Ito & Mester (1999) argue that the lexicon is organized in a similar fashion, and that lexical entries can be classified based on how prototypical or peripheral they are with respect to the phonological constraints of the language. In particular, the Yamato stratum forms the core of the Japanese lexicon, because it obeys the most constraints, while the Foreign stratum forms the outermost periphery of the lexicon because it obeys the fewest constraints (basically, only constraints on syllable structure which are obeyed in all strata). Sino-Japanese falls in between these two layers. The pattern of violations of three of the constraints from Chapter 1, *P, *NT, and CODACOND, illustrate this core-periphery structure. In Yamato *P, *NT, and CODACOND (coda consonants may not have place features—Kager 1999: 131) are all satisfied, in Sino-Japanese only *P and CODACOND are satisfied, and in Foreign only CODACOND is satisfied (Table 3.2). The strata are arranged in a subset relationship, with Sino-Japanese forming a subset of possible Foreign words, and Yamato forming a subset of possible Sino-Japanese words (Figure 3.3). This happens to mirror the etymological history of the different strata—Yamato items having always been a part of the language, Sino-Japanese items having been borrowed relatively long ago, before the 13th c. C.E., and Foreign items having been borrowed relatively recently, from the 16th c. C.E. onwards.[8]

[8] However, note that the vast majority of Sino-Japanese and Foreign items have been borrowed or coined in the past 150 years.

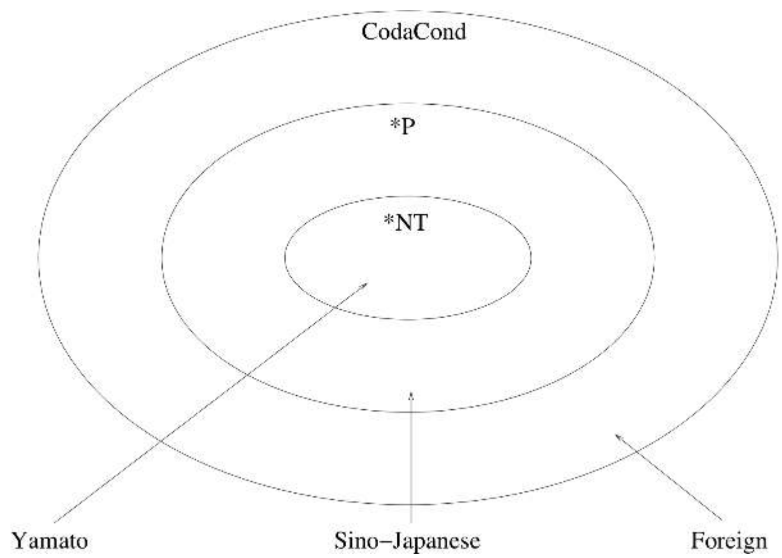


Figure 3.3: The Core-Periphery model of the lexicon

Ito & Mester argue that this core-periphery structure reflects the different degrees of nativization that the different strata have undergone, Foreign items having been nativized less than Sino-Japanese items because they were borrowed later. The process of nativization gradually reduces the number of constraints that a given word violates, causing the word to move further inward, closer to the core of the lexicon. For example, words like 本 /hoɴ/ 'book' and カルタ /karɯta/ 'playing cards', although originally Sino-Japanese and Foreign, respectively, are apparently now fully nativized and have moved into the Yamato stratum, as shown by their participation in native phonological processes like *rendaku* (Chapter 1).

*NT and *P ranking problems

One problem with this model is that the facts are less clear when the Mimetic stratum is taken into account (Inkelas & Zoll 2007). Recall that Mimetic words can violate *P, but not *NT (Table 3.3). Although both Sino-Japanese and Mimetic items form a subset of

Table 3.3: Pattern of violations for *P and *NT, including Mimetic

	Yamato	Sino-Japanese	Mimetic	Foreign
CODACOND	√	√	√	√
*P	√	√	violable	violable
*NT	√	violable	√	violable

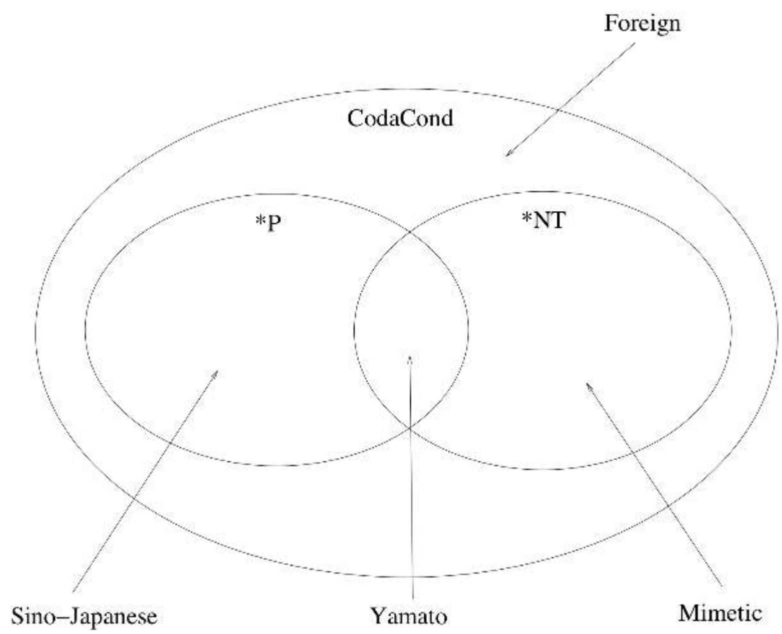


Figure 3.4: The Core-Periphery model of the lexicon (with Mimetic added)

possible Foreign words, it is not possible to place Sino-Japanese and Mimetic in a subset relationship with each other, because there are some Sino-Japanese items which violate *NT and hence cannot be possible Mimetic words, such as 勉強 /beɴkʲoː/ 'study', and likewise there are some Mimetic items which violate *P and cannot be possible Sino-Japanese words, such as ピカピカ /pikapika/ 'sparkle' (Figure 3.4).

Even if the relationship between Sino-Japanese and Mimetic is unclear, it could be maintained that Yamato still serves in some sense as the core of the lexicon, because it is more constrained in terms of possible phonotactic patterns than any of the other strata, at least with respect to the constraints presented so far. Yet there do exist other constraints which govern only Sino-Japanese or Mimetic, putting this view further in jeopardy. While Yamato roots are unrestricted in length, Sino-Japanese roots can only be one or two syllables long (Tateishi 1990), while Mimetic words are always formed from bimoraic (or trimoraic) roots (Poser 1990). The existence of these constraints is unaccounted for in the Core-Periphery model. Yamato, being at the core of the lexicon, should be governed by all of the constraints which govern the other strata, yet these length restrictions on Sino-Japanese and Mimetic do not apply to Yamato, as shown by the existence of one-mora Yamato words, like 木 /ki/ 'tree; wood', and Yamato words which are more than two syllables long, like 紫 /mɯrasaki/ 'purple'.

In fact, it is not even clear that Yamato is really at the core of the lexicon, in the sense of being the most highly constrained stratum. Kawahara, Nishimura & Ono (2003) argue that, although etymologically Yamato forms the core vocabulary of Japanese, in the synchronic grammar of Modern Japanese it is the Sino-Japanese stratum, not the Yamato stratum, which is the most highly constrained. Kawahara et al. mention the length restriction on Sino-Japanese roots, namely that they can only be one or two syllables long, and they also point out that in the case of two-syllable roots, the second syllable can only be /ki/, /kɯ/, /ti/ (=[ʧi]), or /tɯ/ (=[tsɯ]).[9] Yamato morphemes are unrestricted in length or in segmental inventory for the second syllable, however. Also, they note that in noun-noun compounds in which the second noun has penultimate accent, the accent of the entire compound will not fall on the second noun if it is Sino-Japanese, but instead will be shifted back to the first noun. So 席 /séki/ 'seat' normally has an accented first syllable, but not in compounds like 予約席 /jojakɯ́seki/ 'reserved seat', where the accent is placed on the final syllable of /jojakɯ/ instead. Kawahara et al. propose that this is due to a constraint NONFIN[ϕ] which does not allow an accent to fall on the final

[9]Tateishi (1990) attempts to derive all cases of two-syllable Sino-Japanese roots from monosyllabic underlying forms, by analyzing the vowel in the second syllable as epenthetic, only surfacing when necessary to prevent the form from violating CODACOND.

foot of a word. This constraint does not apply to Yamato items, however, as can be seen from compounds like ペルシャ猫 /perɯʃanéko/ 'Persian cat', where the Yamato word 猫 /néko/ 'cat' retains its accent on the first syllable.

Although Ito & Mester (1999) give *NT as an example of a constraint that applies only to the Yamato stratum and not to Sino-Japanese, Kawahara et al. (2003) suggest that *NT is actually obeyed in both strata, but there is a more highly ranked constraint IDENT[VOI](SJ:STEM-INITIAL-σ) (preserve the voicing of stem-initial segments in Sino-Japanese roots) which masks the effect of *NT in Sino-Japanese words. Although this constraint seems rather ad-hoc, Kawahara et al. argue that it is necessary to explain the restrictions on possible segments in the second syllable of Sino-Japanese roots, and that a similar constraint is needed to explain *NT effects in the Mimetic stratum as well. *NT is obeyed within Mimetic roots, but across root boundaries it can be violated, as shown by the reduplicated form トントン /toNton/ 'knock-knock', where the second /t/ cannot be voiced (*toNdon).

Although they are not explicit about this point, Kawahara et al. (2003) seem to be agreeing with Ito & Mester (1999) in assuming that the strata must be placed in a strict subset relation with each other. Otherwise they would not go to the trouble of explaining how *NT could appear to apply only to Yamato and not to Sino-Japanese. Yet they are still unable to explain how there could exist constraints which apply only to a particular peripheral layer, but not to the layers underneath. The length constraint on Mimetic roots is an example of this type of constraint. As I noted before, Mimetic roots must be at least two moras long. Since Mimetic is not at the core of the lexicon, adding this constraint to Kawahara et al.'s (2003) model (Figure 3.5) would predict that Sino-Japanese roots (and possibly Yamato as well, depending on the position of Mimetic with respect to Yamato) are subject to this constraint as well, but this is incorrect given that

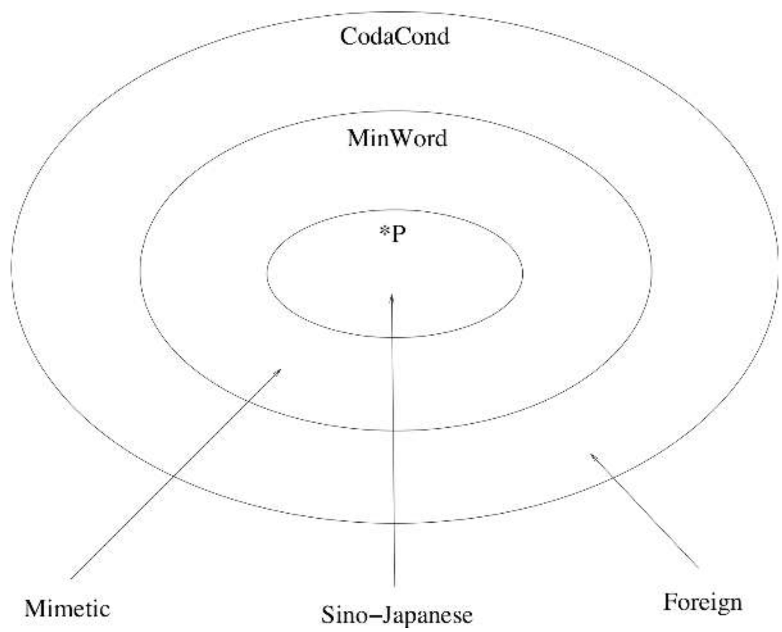


Figure 3.5: The minimum length constraint on Mimetic roots, with Sino-Japanese at the core of the lexicon. This arrangement of the strata incorrectly predicts that MINWORD governs the Sino-Japanese stratum.

there are many examples of one-mora words in both Sino-Japanese and Yamato.

Core and peripheral lexical items

In an earlier account of the Core-Periphery model, Ito & Mester (1995) recognize the existence of stratum-specific constraints like the length constraints governing Sino-Japanese and Mimetic, saying, "...it is not in general possible to impose a total ordering on vocabulary strata..." (p. 820), yet they still maintain that the lexicon as a whole shows an overall core-periphery organization. In addition to *NT and *P, they list several other phonological constraints and processes active in various lexical strata in Japanese (Table 3.4). These constraints interact with each other according to the Venn diagram in Figure 3.6. Ito & Mester (1995) argue that lexical strata result from the overlapping domains of these constraints; for example, they define Yamato as the domain in which *rendaku*, Lyman's Law, *NT, *#R, and *P all hold true. The Foreign stratum, on the other hand, does not have any unique constraints distinguishing it from the other strata, but rather is defined as the region in which the inner constraints (*P, *NT, and *#R) don't apply. Of course, there are plenty of examples of loanwords which do not violate *P, *NT, and *#R, such as タウン /taɯɴ/ 'town'. However, they are not subject to *rendaku* or Lyman's Law, which can be seen in compounds such as ベッドタウン /beddotaɯɴ/ 'bed town' or ニュータウン /nyɯːtaɯɴ/ 'new town', where the initial /t/ in /taɯɴ/ remains voiceless. The Foreign stratum is thus seen as a region of the lexicon where any phonotactic pattern can occur, subject only to the syllable structure constraints like CODACOND that govern all strata:

> ...the large and very heterogeneous class of Foreign items...should not be considered as constituting a uniform stratum. Rather, we are simply dealing with less central areas of the lexicon, where more and more constraints are violated. (p. 824)

Rice (1997) criticizes Ito & Mester's (1995) approach, pointing out that obeying a particular constraint is a necessary, but not sufficient, condition for classifying a given word. For example, Sino-Japanese roots are mono- or bisyllabic, but the converse does not hold—not all monosyllabic roots are Sino-Japanese. So given a monosyllabic word like 木 /ki/ 'tree; wood', it is not possible to tell from just the phonological form of the word which stratum it belongs in. It could potentially belong to either Yamato or Sino-Japanese.[10] In this case we would need to look for other types of evidence to determine the stratum which /ki/ belongs in. In particular, since /ki/ undergoes *rendaku* in compounds such as 柏木 /kaʃiwa+ki/ → [kaʃiwag̲i] 'oak tree', this suggests that /ki/ belongs in the Yamato stratum.

[10] But not Mimetic, because Mimetic roots are always bimoraic. Also, as I will show later, this cannot be a possible Foreign word because Foreign words must be a minimum of two moras long.

Table 3.4: Phonological constraints distinguishing lexical strata, from Ito & Mester (1995)

RENDAKU	Word-initial obstruents in second compound member must be voiced
LYMAN'S LAW	No more than one voiced obstruent per morpheme
ROOT= σ	Root is exactly one syllable long
ROOT= ϕ	Root is exactly one foot long (= two moras)
*NT	Voiceless stops cannot occur after /ɴ/
*P	No single (i.e. not geminated or post-nasal) /p/
*#R	No word-initial /ɾ/
*ČE	Palatal consonants cannot occur before /e/
*TI	Non-palatal coronals cannot occur before /i/
*DD	No voiced geminates
*F	/ɸ/ only occurs before /ɯ/
*TS	/ts/ only occurs before /ɯ/
□FU	/h/ cannot occur before /ɯ/
□TSU	/t/ cannot occur before /ɯ/
*YE	/j/ cannot occur before /e/
*SI	/s/ or /z/ cannot occur before /i/

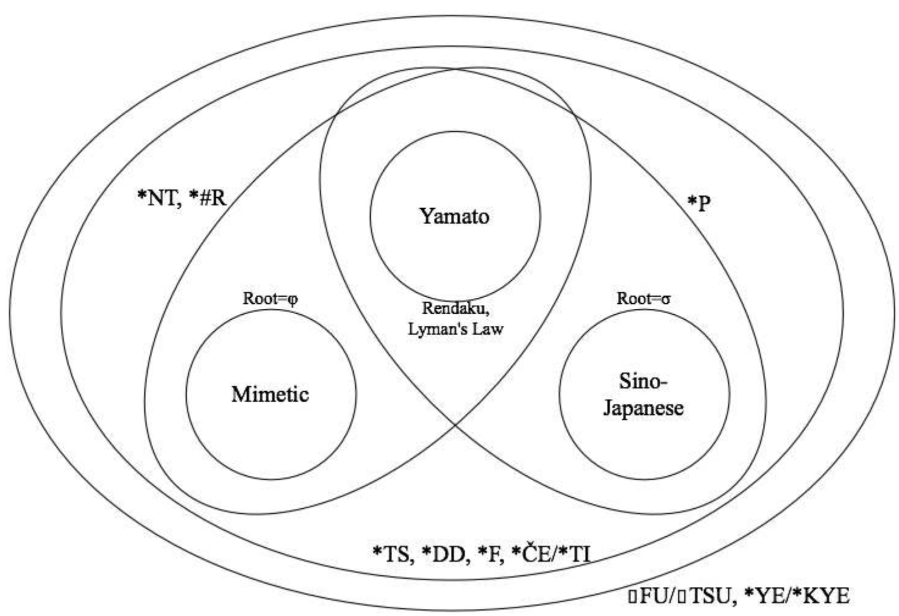


Figure 3.6: Domains of phonological constraints, from Ito & Mester (1995)

Table 3.5: Yamato *NT violations, from Rice (1997)

いんちき	iɴʧiki	‘trickery’
あんた	aɴta	‘you’ (< あなた /anata/)
ちゃんこ鍋	ʧaɴkonabe	‘sumo-wrestler stew’

Rice also notes that the *NT constraint is not absolutely inviolable in the Yamato stratum, as shown by the examples in Table 3.5. In these cases we have words whose etymology places them in the Yamato stratum, yet they each contain a post-nasal voiceless obstruent. あんた /aɴta/ ‘you’ is a particularly interesting case of a *NT exception, because it coexists with あなた /anata/, the form from which it is derived via syncope. Although Ito & Mester (1995) do not address the specific case of /aɴta/, they do note the existence of historically native words like ちぇっ /ʧeꞯ/[11] (a swearword) and しぇえ /ʃeː/ (an exclamation used by a famous cartoon character) which violate another of their proposed markedness constraints, *ČE,[12] saying,

> It is also important not to entirely equate “peripheral” with “foreign.” Violations of the *ČE-constraint are not restricted to recent loans, but are also found among items of native origin.... Such forms are undoubtedly native, but peripheral. (p. 830)

Ito & Mester are forced to conclude that words like /aɴta/ are peripheral forms because they define strata solely in phonological terms, but it is not clear in their model what part of the lexicon /aɴta/ would be located in. It’s not a possible Mimetic root, because it violates *NT. It could be a possible Sino-Japanese word, but this would have to involve some sort of folk etymology where /aɴta/ was composed of two Sino-Japanese roots

[11] Here /ꞯ/ is realized phonetically as a glottal stop [ʔ] (Vance 1985, 1987).

[12] Another example of a native word that violates *ČE is ひぇえ /h^{j}eː/, which Akamatsu (1997) glosses as “an interjection expressing great fear” (p. 90). Note that all of these *ČE-violating examples happen to be interjections, which corroborates their peripheral status in the Yamato stratum.

/aɴ/ and /ta/; /aɴta/ itself is not a possible bisyllabic Sino-Japanese root because the final syllable is not one of /ki, kɯ, ʧi, ʦɯ/. But this seems extremely unlikely, since the existence of both /aɴta/ and /anata/ in common usage makes it hard to believe that Japanese speakers would not know that these two words are related.

Constraints on the Foreign stratum

Further evidence against the Core-Periphery model can be found by looking for constraints which apply to loanwords, but which are freely violated in one or more of the other strata. In particular, there seems to be a minimal length constraint on possible Foreign words which does not govern the Sino-Japanese and Yamato strata. Both Lovins (1975) and Tsuchida (1995) note that the tense-lax distinction in English vowels is usually represented by length in the corresponding borrowings in Japanese: tense vowels are borrowed as long (bimoraic) vowels, while lax vowels are borrowed as short (monomoraic) vowels. Lax vowels can only occur in closed syllables in English, and since Japanese does not allow coda consonants in general (with the exception of moraic nasals, and the first segment of a geminate), a coda consonant in the source word will have an epenthetic vowel (usually /ɯ/) inserted after it.[13] This results in every loanword from English having at least two moras.

Some examples of loanwords derived from monosyllabic English words are given in Table 3.6. In every case, the loanword is at least two moras long. For the source words with tense vowels this follows directly from the fact that tense vowels are borrowed into Japanese as long vowels. For the source words with lax vowels, which are always closed syllables in English, the second mora comes from the epenthetic vowel inserted after the coda consonant.

[13] Under certain conditions, the coda consonant will be geminated as well, as discussed in section 2.3.4 of Chapter 2.

Table 3.6: Adaptation patterns of English monosyllables

Open σ, tense V	Closed σ, lax V	Closed σ, tense V
key → kiː	*lip* → rippɯ	*cheap* → ʧiːpɯ
pay → peː	*pet* → petto	*cape* → keːpɯ
show → ʃoː	*loss* → rosɯ	*rope* → roːpɯ
you → jɯː	*put* → pɯtto	*hoop* → ɸɯːpɯ

Table 3.7: Examples of truncated loanwords

2 moras	k^jara	←	k^jarakɯtaː	'character'
	ana	←	anaunsaː	'announcer'
	ʧoko	←	ʧokoreːto	'chocolate'
3 moras	arɯmi	←	arɯmin^{j}ɯːmɯ	'aluminum'
	koɴbi	←	koɴbineːʃoɴ	'combination'
	terebi	←	terebiʤoɴ	'television'
4 moras	pasokoɴ	←	paːsonarɯ+koɴp^{j}ɯːtaː	'personal computer'
	apaːto	←	apaːtomeɴto	'apartment'
	koɴsaba	←	koɴsaːbatibɯ	'conservative'

A similar pattern arises in truncated versions of loanwords (Table 3.7). Usually these are formed from the first two moras of the loanword. Although there are also some examples that are three or four moras long, there are never any that are only one mora long (Ito 1990, Labrune 2002). Both of these facts suggest that there is some kind of minimal length constraint MINWORD (a word must have at least two moras) operating on words in the Foreign stratum.[14]

There are only three loanwords listed in JMDICT (Electronic Dictionary Research and Development Group 2003) which violate MINWORD. One of them, ザ /za/ 'the', although listed as a separate word, actually only occurs as a bound morpheme in titles and such, so it is not a true exception. The other two exceptions, ティ /ti/ 'tea' and ディ

[14]For ease of exposition I have formulated this as a single constraint, but Kager (1999) suggests that minimal length effects actually arise from the interaction of two constraints, FTBIN (feet are binary in terms of syllables or moras) and GRWD=PRWD (a grammatical word is a prosodic word).

/di/ 'day', have two-mora variant pronunciations, ティー /tiː/ and デー /deː/, which are more common (John Whitman, p.c.). As well, Lovins (1975) notes that one-mora loans are short-lived and quickly replaced by two-mora variants created by lengthening the vowel, which is probably what is happening to /ti/ and /di/.

However, while one-mora words are not possible in the Foreign stratum, they do exist in the Sino-Japanese and Yamato strata, and include such common words as 木 /ki/ 'tree; wood', 手 /te/ 'hand', 日 /hi/ 'sun; sunshine; day', and 目 /me/ 'eye'. Thus MINWORD is an example of a constraint which applies in the Foreign stratum, but not in Yamato or Sino-Japanese, and is further evidence against there being an overall core-periphery organization to the lexicon.

Thus, the Core-Periphery model does not provide an adequate model of the relationships between the Yamato, Sino-Japanese, and Mimetic strata. Yamato cannot be defined solely as the intersection of the domains of *NT, *P, and other markedness constraints, because there are a small number of exceptions, like /aNta/, which belong in Yamato yet violate *NT. As well, Yamato, Sino-Japanese, and Mimetic cannot be placed in a subset relationship with each other, because Sino-Japanese, Mimetic, and Foreign all have different length constraints on possible words, none of which are obeyed in Yamato.

3.3.2 The Cophonology model

An alternative to the Core-Periphery model is the *Cophonology* model (Anttila & Cho 1998, Anttila 2002, Inkelas & Zoll 2007), which allows for different constraint rankings for different classes of words or morphemes. In fact, the model proposed in Ito & Mester (1995) is itself a kind of cophonology model, since it recognizes that some con-

Table 3.8: Constraint rankings for Japanese lexical strata

Yamato	*NT, *P ≫ FAITH
Sino-Japanese	*P ≫ FAITH ≫ *NT
Mimetic	*NT ≫ FAITH ≫ *P
Foreign	FAITH ≫ *NT, *P

Table 3.9: Constraint rankings for assimilated and unassimilated loanwords

Yamato, Sino-Japanese, Mimetic	CODACOND, *TI, *DD, *NT, *P ≫ FAITH
Assimilated Foreign	CODACOND, *TI, *DD ≫ FAITH ≫ *NT, *P
Unassimilated Foreign	CODACOND ≫ FAITH ≫ *TI, *DD, *NT, *P

straints like *P apply only in some strata and not in others; however, Ito & Mester still say that the overall organization is that of a core-periphery structure, and in later work have abandoned their 1995 model entirely in favor of a strict core-periphery model. A Cophonology analysis of the Japanese lexicon would allow for Yamato, Sino-Japanese, and Mimetic to have differing relative rankings for the markedness constraints *P and *NT with respect to faithfulness, as shown in Table 3.8. Here we can see that FAITH can be freely ranked with respect to the other two constraints, and each possible ranking corresponds to a specific lexical stratum. The relative ranking of FAITH with markedness constraints like *TI and *DD which are obeyed in the non-Foreign strata also allows us to distinguish between *assimilated* loanwords (like ジレンマ /ʤireɴma/ 'dilemma') which obey these constraints, and *unassimilated* loanwords (like パーティー /paːtiː/ 'party') which do not (Table 3.9).

Following Anttila & Cho (1998) and Anttila (2002), the constraint rankings in Tables 3.8 and 3.9 can be organized in a *grammar lattice*, which is a tree-like structure in which the topmost node contains the most general rankings applicable to all strata, and lower nodes containing progressively more specific rankings narrowing down to the specific strata themselves. A grammar lattice for Japanese is shown in Figure 3.7.

While the Cophonology model of lexical stratification would be able to adequately

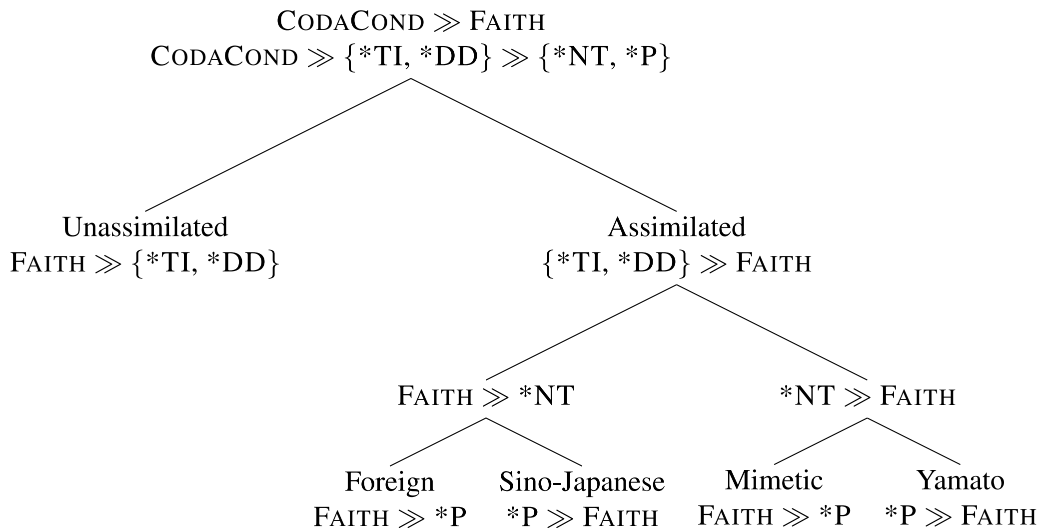


Figure 3.7: Grammar lattice for Japanese lexical strata

account for the Japanese facts, it is conceptually more complex, and therefore less predictive, than the Core-Periphery model. In the Core-Periphery model, the overall constraint ranking is primary, while lexical strata are an emergent property based on this ranking (so that Yamato is defined as the set of items in which faithfulness is ranked below all markedness constraints, while Sino-Japanese is defined as those items for which faithfulness is ranked above *NT but below *P, and so on). In the Cophonology model, on the other hand, both the individual strata and the constraint rankings for each stratum must be stipulated, since the constraint rankings can vary arbitrarily across the strata. Because of this, it is difficult under the Cophonology model to explain why lexical items like 本 /hoɴ/ ‘book’ and カルタ /karɯta/ ‘playing cards’ (both of which undergo *rendaku* despite being etymologically Sino-Japanese and Foreign, respectively) have changed their stratal assignment over time. Under the Core-Periphery model, stratal reassignment has a natural explanation: as a word becomes more nativized over time, it gradually falls under the domain of more and more markedness constraints, causing it to move closer to the core of the lexicon. In the Cophonology model, on the other hand, there is no single stratum which is distinguished as the core of the lexicon, and so it is less obvious why nativization would cause lexical items to be reclassified at all over time, not to mention why they are reclassified as Yamato instead of one of the other strata.

3.4 The representation of phonological neighborhoods

The problem with both the Core-Periphery and Cophonology models of lexical stratification is that neither can account for the type frequency and phonological neighborhood effects which influence how adaptation patterns change over time. Phonological constraints such as *TI are too coarse-grained to pick out specific environments such as word-final /−ti/, and thus cannot account for the behavior of speakers in the late 19^{th} century, who tended to preserve /t/ in this environment in loanwords while palatalizing in other environments. In addition, neither model directly represents either type or token frequency, and thus cannot easily explain why non-native patterns occurring in phonological neighborhoods with a higher type frequency were less likely over time to be nativized.

The sensitivity of adaptation patterns to the phonological neighborhood in which they occur in may have an explanation in terms of the effects of lexical neighbors on the perception and production of a lexical item (Luce & Pisoni 1998, Vitevitch & Luce 1999). A lexical neighbor of a word is another word which differs by only a single phoneme, so for example the lexical neighbors of *cat* are *rat*, *cot*, *cap*, *cast*, *at*, and so on. (In OT terms, the lexical neighborhood of a word would be the set of forms which would entail only a single violation of faithfulness.) Dell & Gordon (2003), in reviewing the literature on lexical neighborhood effects, state that words with higher-density lexical neighborhoods take longer to identify, and are identified with lower accuracy, than words with low-density neighborhoods. This is because, during the perception of a word, the lexical neighbors are partially activated in the lexicon as well, making them competitors with the target word. Thus, the more lexical neighbors a word has, the harder it is to identify it correctly. On the other hand, a dense lexical neighborhood has the opposite effect on production. Having many lexical neighbors makes it easier to produce a target word. Dense lexical neighborhoods also improve sublexical processing (Stemberger 2004). These facts suggest that as more loanwords containing a non-native pattern such as TI are learned by a speaker, it would become easier to both perceive and produce the contrast between the non-native and nativized pattern (such as the difference between /ti/ and /ʧi/).

In order to explain how these kinds of neighborhood effects arise in lexical processing, a model of the lexicon is needed in which lexical entries are organized by frequency and phonological similarity, such that the processing of a word can be influenced by other words which are sufficiently similar to it. This is the assumption made in many of the lexical models proposed in psycholinguistic research, such as the TRACE model (McClelland & Elman 1986), as well as in exemplar-based models (Pierrehumbert 2001). One particularly promising model is the distributed, interactive processing framework for word reading proposed in Seidenberg & McClelland (1989). Connectionist implementations of this framework (Plaut, McClelland, Seidenberg & Patterson 1996, Harm & Seidenberg 1999) have been developed as an alternative to traditional *dual-route* models of reading (Coltheart 1981, Coltheart, Curtis, Atkins & Haller 1993). In a dual-route model, there are separate mechanisms handling the regular and exceptional aspects of language processing. For example, a dual-route model of past tense morphology in English would have both a rule which adds −*ed* to a verb stem to form the regular past tense (e.g. *walk*→*walked*), and a list of exceptions to which the rule does not apply (*run*→*ran*; Pinker 1991). This corresponds to the traditional division in linguistics between the grammar and the lexicon, with the grammar being the set of rules that apply to regular forms, and the lexicon the list of exceptional forms. In a *single-route* model, on the other hand, both regular and irregular verbs would be handled by the same mechanism, and thus there is no clear division between the grammar and the lexicon. Single-route models can involve either a full listing of all pairs of non-past and past tense forms for all verbs in English (redundantly listing even the regular forms which can be derived by rule), as in exemplar models, or conversely a set of rules for inflecting the past tense which vary in generality (Stockall & Marantz 2006). Unlike dual-route models, both kinds of single-route models are able to capture the fact that the exceptional items in processes like past tense formation often have subregularities of their own, such as groups of verbs which undergo the same kind of vowel alternation (*sing*→*sang*, *drink*→*drank*, and so on; Bybee & Slobin 1982).

Seidenberg (2005) discusses this *quasiregular* nature of many linguistic generalizations, pointing out that " '[r]ule-governed' forms and 'exceptions' represent points on a continuum of spelling-sound consistency. Many aspects of language have this graded character." (p. 239) The connectionist models of Seidenberg & McClelland (1989) and Plaut et al. (1996) represent these quasiregularities directly by learning mappings from orthography to phonology with varying degrees of generality. For example, a typical network in one of these models learns to read an initial ⟨b−⟩ as /b−/, and discovers during the course of learning that this mapping is extremely general, as there are no exceptions encountered in the training data. On the other hand, the mapping ⟨−ave⟩→/−eɪv/ is almost as general as the mapping ⟨b−⟩→/b−/, however there are exceptions such as ⟨have⟩→/hæv/. The fact that the Seidenberg & McClelland (1989) and Plaut et al. (1996) models of word reading can replicate the cross-cutting effects of word frequency and orthography-to-phoneme consistency on word reading times and error rates seen in human subjects shows that these models, in which the lexicon is represented as a set of mappings between phonological, semantic, and orthographic representations that are sensitive to similarity and frequency, can be plausible models for how lexical items are actually represented and processed in the brain.

Of course, these quasiregular aspects of language processing can be replicated to some degree in generative models of phonology as well, through such mechanisms as rule ordering or constraint ranking. However, generative models do not typically have a causal role for word or type frequency, or for phonological similarity. For these properties to be able to affect lexical processing in a generative model, it would have to be stipulated that certain rules or constraints are more likely to apply to high-frequency items, or to items in dense lexical neighborhoods, and so forth. Yet these kinds of similarity and frequency effects arise naturally in connectionist models, due to the ways in which connectionist networks process their inputs and learn how to perform language tasks:

> The twin principles of learning and distributed representations enable connectionist models to explain frequency effects and some similarity effects 'for free'. Learning is, by its very nature, sensitive to frequency. The more a unit in the network is activated, the more that connections involving it grow stronger. And the learning that occurs for a particular item automatically applies to similar items because similar items share units and connections. (Dell 2000: 346)

Thus, a connectionist model of loanword adaptation may provide a way to explain the frequency and phonological neighborhood effects seen in loanword adaptations, as well as the tendency for some of these adaptation patterns to be generalized to larger and larger phonological neighborhoods over time, in a process resembling the lexical diffusion of sound change (Wang 1969, Labov 1981). In particular, if we consider adaptation strategies like TI→TI or TI→ČI to be mappings of varying regularity from an L2 source to an L1 phonological representation, then they seem to have similar properties to Seidenberg's (2005) notion of quasiregularities. Adaptation conventions, like other kinds of quasiregularities in language processing, can range from the highly general, such as the context-free mapping of English stops to their corresponding Japanese stops, to less general patterns, such as the adaptation of velars occurring before /æ/ in English as palatalized velars in Japanese, down to highly specific patterns, such as the historical adaptation of English /t/ occurring before a word-final /i/ or /ɪ/ as Japanese [t] or [ʧ].

In the next chapter, I develop these ideas further with the design and analysis of two connectionist networks, one which is trained to identify the lexical stratum of words presented to it, and the other which is trained to perform loanword adaptations.

3.5 Conclusion

In this chapter I have examined the process of adaptation of loanwords by a single speaker. While prior analyses have focused either on the segmental parse (SP) stage or the constraint satisfaction (CS) stage in explaining adaptation patterns, I have argued that both must be considered to give a complete account of loanword adaptation that can be applied to any language contact situation. Likewise, the debate over whether the input to borrowing is phonetic or phonological in nature is similarly misguided, since either may be possible depending on the type of contact taking place. In fact, most of the cases of borrowing in the literature that are claimed to be examples of phonological adaptation can be argued to actually be examples of orthographically-based adaptation. This is definitely true of English borrowings into Japanese, for which the majority enter the language through the written route, not the spoken route. While there is evidence for both spoken and written adaptations in historical borrowings, written borrowings predominate among recent loanwords.

I then considered what the historical development of adaptation patterns can tell us about the organization of the Japanese mental lexicon. The way in which adaptation patterns change over time, with foreign phonotactic patterns first being attested in high-frequency phonological neighborhoods, then later spreading to other, lower-frequency environments, suggests that the representation of loanwords in the mental lexicon is sensitive to frequency and phonological similarity effects. These kinds of effects arise naturally in connectionist models based on the Seidenberg & McClelland (1989) framework for lexical processing, suggesting that the development of a connectionist model of loanword adaptation may be more insightful than the Core-Periphery or Cophonology models of the lexicon, both of which classify lexical items into groupings (lexical strata) that are too coarse-grained to be able to account for similarity and frequency effects on adaptation patterns over time.

In the next chapter, I will flesh out this idea by presenting an implementation of such a model. I will then use this to model to examine the adaptation patterns governing voiced geminates presented in Chapter 2, showing that the relatively acceptability of /dd/ and /gg/ as compared to /bb/ in loanword derives at least in part from the different frequencies of the voiceless counterparts /tt/, /kk/, and /pp/ in the Yamato and Sino-Japanese strata of the Japanese lexicon.

CHAPTER 4

A CONNECTIONIST MODEL OF LOANWORD ADAPTATION

In the previous chapter I examined two competing theories of lexical stratification: the Core-Periphery model (Ito & Mester 1999), and the Cophonology model (Anttila 2002). The Core-Periphery model is too restrictive to provide an adequate account of lexical strata in Japanese, because there are different minimal length constraints governing the four strata which cannot be ordered in a single ranking. The Cophonology model, on the other hand, is not restrictive enough, since it allows for potentially arbitrary differences in constraint rankings between the four strata, and thus cannot explain how lexical entries like 本 /hoɴ/ 'book' have been reassigned to the Yamato stratum over time. In addition, neither model can be used to explain how type frequency and phonological similarity can cause changes in adaptation patterns over time, since there is no account in either framework for how lexical entries are organized with respect to phonological similarity, nor for how phonological constraints like *TI are inferred from the distribution of phonotactic sequences in the lexicon.

In this chapter I develop an alternative model of the Japanese lexicon, based on connectionist models of single word reading such as Harm & Seidenberg (1999). Connectionist models introduce a similarity metric for lexical representations, which allows for different constraint rankings to govern different subsets of the lexicon (as in the Cophonology and Core-Periphery models), as well as for the processing of a lexical item to be influenced by similar items in the lexicon, providing a possible mechanism for both loanword nativization and diachronic changes in adaptation patterns. I begin with a brief introduction to connectionism, highlighting the differences between OT and connectionist theories of phonology and the lexicon, and introducing Seidenberg & McClelland's (1989) general framework for connectionist lexical representation and processing. I then present an analysis of a connectionist network trained to classify lexical entries by stratum (Yamato, Sino-Japanese, or Foreign), and use the results from this network to explore how Japanese lexical strata are structured as clusters of phonological neighborhoods. Finally, I develop a connectionist model of loanword adaptation, and examine the effects of lexical frequency and phonological similarity on the adaptation of voiced geminates in loanwords.

4.1 Connectionist models of cognition

Connectionist (or "neural network") models are a class of statistical models which are inspired by—though not necessarily a faithful model of—the cellular organization of nervous systems.[1] In a connectionist network (CN), there are a number of simple processing units, and a set of weighted connections between these units. Each unit u_i has an associated *activation function* f_i, which determines the activation a_i of the unit based on the weighted sum of the activations a_j of all of the units feeding into it:

$$a_i = f_i(\sum_j w_{ij} a_j + \theta_i). \tag{4.1}$$

(The θ_i term in this equation is known as a *bias*, and helps determine the unit's threshold of activation.) The activation function f_i is often a sigmoid (or "S-shaped") curve, such as the logistic function:

$$f(x) = \frac{1}{1+e^{-x}}, \tag{4.2}$$

or the hyperbolic tangent:

$$f(x) = \frac{e^{2x}-1}{e^{2x}+1}. \tag{4.3}$$

Graphs of these two functions are shown in Figure 4.1. Using a sigmoid as the activation

[1] Bishop (1995), Ripley (1996), and Sarle (2002) discuss the relationships between connectionist models and other statistical models. In particular, Ripley (1996) notes that connectionist models serve as a useful bridge between highly-constrained, yet inflexible, statistical models with a small number of parameters, and completely unconstrained, non-parametric models.

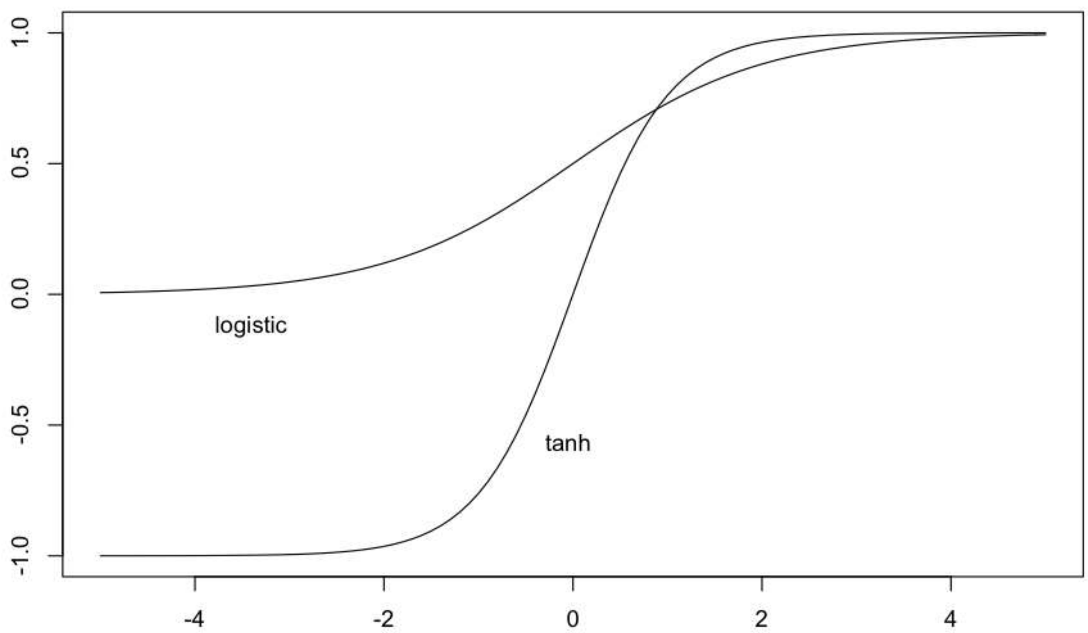


Figure 4.1: The logistic and tanh functions

function for a unit introduces a nonlinearity in the unit's response. A unit's output activation does not linearly increase with increasing input activation, but instead grows asymptotically closer to its maximum value (which is 1 for both the logistic and tanh functions). Likewise, as input activation decreases, output activation approaches the minimum value for the activation function (0 for logistic, -1 for tanh).

The computational power of a CN comes from linking the relatively simple processing units together in various configurations, allowing the network as a whole to be able to compute more complex functions of its input than a single unit on its own would be capable of. The units in a CN are organized into groups called *layers*, which can be thought of as corresponding in some ways to different levels of representation in linguistic theories. A common architecture for a CN is known as a *feedforward* architecture, as shown in Figure 4.2. In this type of network, there is a single *input* layer of units, where an input representation is presented to the network; an *output* layer, where the network's response to the input is represented; and usually one or more layers of *hidden*

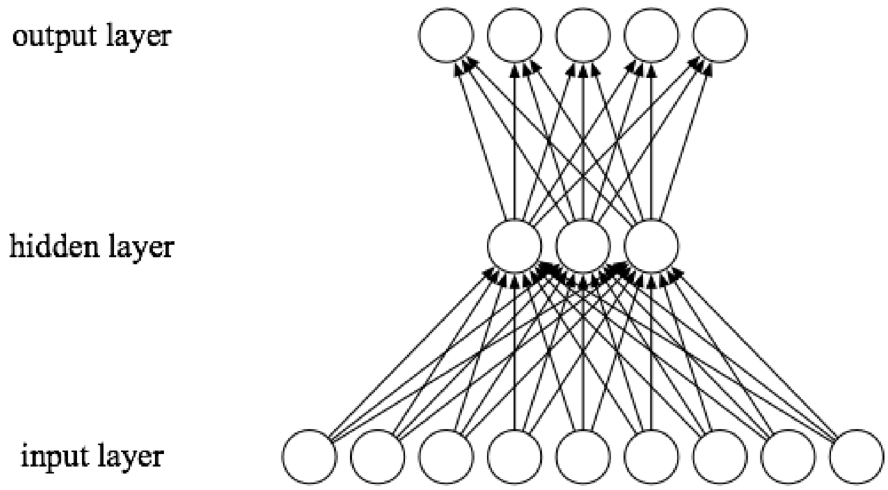


Figure 4.2: A feedforward connectionist network. The circles represent processing units, while the lines represent the weighted connections between units. Activation flows from the input layer upwards to the output layer.

units, which mediate between the input and output layers, and which allow the network to form internal representations of the input data.[2] Each of the units in the input layer are connected to every one of the units in the first hidden layer, then each one of these units in the hidden layer are connected to the units in the next layer, and so on until the output layer is reached. There are no connections from a higher-level layer to a lower-level one; instead, the unit activations in each layer depend only on the unit activations in the previous layer. An input is processed in the network by setting the activations of the input layer units, then allowing each successive layer to compute their unit activations based on the weighted sum of the units in the previous layer, until finally the flow of activation in the network reaches the output layer. A feedforward network can be seen as computing a function on its input, and it has been shown that a single hidden layer is sufficient for these types of networks to be able to approximate any continuous function

[2]It is also possible to construct a feedforward network with only an input and output layer, and no hidden layers. Such a network is generally known as a *perceptron*, and in fact were some of the first connectionist models to be investigated (Rosenblatt 1958, 1962). However, perceptrons have interesting limitations on their computational power. For example, they cannot be used to solve the “xor” problem, in which a network with two input units must activate an output unit when only one or the other of the inputs is activated, but not both (Minsky & Papert 1969). A feedforward network must have at least one hidden unit in order to be able to learn this task.

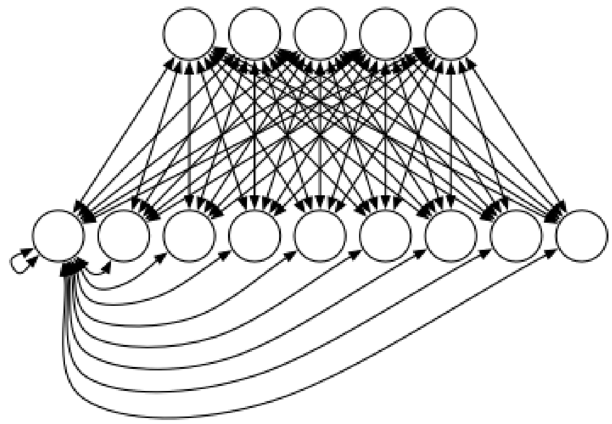


Figure 4.3: An attractor network, with recurrent connections both within the representation layer, and between the representation and cleanup layers. (The layer-internal recurrent connections for only one of the representation units are shown here.) Activation flows back and forth between the representation layer and the cleanup layer for a specified number of time units.

(Hornik, Stinchcombe & White 1989).

Besides the feedforward architecture, there are many other possible ways to organize the units in a CN. Another common network architecture, one which will be used in section 4.5 for the loanword adaptation network, is known as an *attractor* network (Figure 4.3). In this type of network, there is a single layer of units (the *representation* layer) which functions both as an input layer and as an output layer, and which has weighted connections between all of the units in this layer. There is also a second layer of *cleanup* units, with bidirectional connections between the representation and cleanup layers. Processing in this network takes place when the activations of the representation units are set to a (possibly noisy) input, then activation flows back and forth between the representation and cleanup layers for a number of discrete time units, until the activations on the representation layer settle to a stable pattern, which is known as the *attractor* for the input.

Despite the diversity of network architectures possible in connectionism, the underlying principle by which all of these network models produce an output is fundamentally the same in each case. First a set of units is set to have an activation pattern corresponding to a particular input representation. Then the units in the network repeatedly calculate their activations for a specified period of time based on the weighted sum of the activations of the units connected to them. Finally, an output representation is obtained from the activations of a designated set of output units.

Within this very broad framework of connectionist modeling, specific models of language processing can be designed and evaluated, by framing a language task (such as learning to produce the surface phonetic form of a lexical item, given an orthographic representation) in terms of pairs of input-output representations to be learned. The function that the network computes for each input depends on the values of the weights w_{ij} connecting the units in the network, and so the goal during the training of a network is to find a set of weights which not only produces the correct outputs for the input patterns used during training, but also generalizes to new inputs not seen by the network during training. There are a variety of algorithms for finding these weights, depending on the architecture of the network. The most commonly used training algorithm for feedforward networks is known as *backpropagation* (short for "backwards propagation of error"), in which the error (the difference between the target output, and the actual output of the network) and the current values of the weights in the network are used to assign an amount of "blame" to each unit for producing the erroneous output. This is used to compute an error gradient for the current values of the weights, which are then adjusted proportionally to the negative of this gradient so as to reduce the error on future presentations of the current input (Rumelhart, Hinton & Williams 1986, Reed & Marks 1999):

$$\Delta w_{ij} = -\varepsilon \frac{\partial E}{\partial w_{ij}}. \tag{4.4}$$

Here E is the error value for the current input-target pair, and ε is a parameter ranging from 0 to 1 known as the *learning rate*. Generally speaking, the larger the learning rate, the fewer training trials are needed for the network to learn the task correctly; however, too large of a learning rate can make learning unstable and cause the network to not learn the task at all (Reed & Marks 1999).

It is clear from this discussion that the methodology used in connectionist accounts of language processing is quite different from that of generative linguistics. In generative linguistics, we proceed by analyzing a body of language data to find a sufficiently general set of rules or constraints which govern the patterns found in the data, then create a model explicitly embodying these rules or constraints. The focus is on the linguistic knowledge that a speaker must acquire and use in order to perform various language tasks. In connectionism, on the other hand, a specific language task is identified, a set of training data is created, and a network is then trained to perform this task. The focus is instead on creating a system which, in the process of learning how to do a language task, will thereby implicitly learn the rules or constraints governing the language data it is being exposed to. As Dell (2000) puts it:

> Connectionism, in its broadest conception, is not a theory of learning, cognition, or perception. It is a language for expressing such theories. But connectionist principles are much better suited to some kinds of theories than others. The kinds of theories that benefit most from a connectionist perspective are those that emphasize the role of learning and recent experience, graded rules, constraint satisfaction, and how knowledge is used in actual tasks. (Dell 2000: 348)

This is not to say that a CN does not develop any kind of abstract knowledge component at all in the process of learning to perform a specific language task. Hanson & Burr (1990) point out that hidden units can be used by a CN to create internal transformations of the input data; specifically, to "construct *variables* (units that respond consistently to the same input category) and *predicates* (units that respond to a value of a variable with a consistent output) that may be useful in solving a problem that the net has been given" (p. 476). More generally, the debate about whether or not networks have rules (Pinker & Prince 1988) has stemmed to some degree from different ideas over what a rule actually is:

> To say that a network does not have rules is factually incorrect, since networks are function approximators and functions are nothing if not rules. So arguments about whether or not networks have rules really do not make much sense.... What we take as a more interesting question is, *What do the network's rules look like?* Are they merely notational variants of the rules one sees in more traditional approaches such as production systems or linguistic analyses? Or do they make use of primitives (representations and operations) which have significantly different properties than traditional symbolic systems, and which might capture more accurately—and with more explanatory power—the behavior of learning in humans? ... So we believe that connectionist models do indeed implement rules. We just think those rules look very different than traditional symbolic ones. (Elman, Bates, Johnson, Karmiloff-Smith, Parisi & Plunkett 1996: 102–3)

Here lies the appeal of connectionist models for explaining language phenomena in which similarity and frequency play a role, such as the loanword adaptation patterns presented in Chapter 2. In generative models of phonology, similarity and frequency typically have no causal role to play in determining the well-formedness of a given lexical item. There is no way to easily account for these effects, except by stipulating that the likelihood of a particular rule or constraint being used depends on the frequency of the lexical item in question, or on its similarity to another item. CNs, on the other hand, induce a similarity metric on the input representations that they are exposed to, as well as the internal representations that they develop, and this metric may be skewed by the frequency of the input items it is trained on over the course of learning (Plaut et al. 1996: 100). Frequency and similarity effects thus arise naturally from the learning and processing mechanisms of the network, allowing for a more parsimonious account of these effects in language processing and language change.

4.2 Constraint-based theories of phonology and lexical processing

As Prince & Smolensky (1993) acknowledge, Optimality Theory has its origins in connectionist models of language processing, via the theory of Harmonic Grammar (HG; Legendre, Miyata & Smolensky 1990, Pater 2008). Grammars in HG are similar to those in OT in that both consist of a set of constraints with varying degrees of strength in the grammar. The difference is that in HG, the constraints have a numerical weight associated with them, while in OT, constraints have a ranking instead.[3] Mapping from an input to an output form in both OT and HG involve the simultaneous satisfication of a set of constraints on output forms, and a CN as well can be seen as an implementation of a constraint-satisfaction system (Dell 2000). In fact, Smolensky & Legendre (2006) argue that CNs and generative models (in the form of Optimality Theory) are really compatible theories at different levels of abstraction, with OT being a high-level theory of language behavior, connectionism being a low-level theory, and HG being an intermediary-level theory linking the two. An OT grammar can be approximated by an HG grammar with a suitable set of weights for its constraints,[4] while an HG grammar

[3]There are some other versions of OT, such as Stochastic OT (Boersma & Hayes 2001), which also use numerically-weighted constraints.

[4]Specifically, if constraint A outranks constraint B in an OT grammar, then the weight of constraint A in the HG approximation must be greater than the weight of constraint B multiplied by the maximum number of times any form can violate constraint B (Prince & Smolensky 1993).

can be implemented by certain kinds of CNs known as *harmonic networks* (Smolensky 2006b).

Smolensky & Legendre's (2006) hypothesis that connectionist models can be described at a higher level by OT grammars is appealing, since it provides a bridge from the linguistic analysis of phonological patterns down to a connectionist implementation of such patterns, and hopefully from there to an understanding of the actual neural structures underlying human language knowledge and performance. Yet, as I discussed in Chapters 2 and 3, an OT grammar is too abstract to be able to explain the frequency and similarity effects that are exhibited in the attested variation in loanword adaptation patterns. On the other hand, connectionist models of single-word reading, such as those proposed by Plaut et al. (1996) and Harm & Seidenberg (1999), have been quite successful at replicating the effects of word frequency and spelling regularity, and the interaction between the two, on word reading times, suggesting that they may be useful for accounting for loanword variation as well. These models are set within the general framework for lexical processing introduced in Seidenberg & McClelland (1989), commonly known as the "triangle model" (TM; Figure 4.4). In TM, a single word has three levels of representation: an orthographic representation, a phonological representation, and a semantic representation.[5] In an implementation of TM, these levels of representation correspond to distinct layers of processing units, with layers of hidden units mediating between each pair of representation layers. There are two key features of TM. First, representations are *distributed*, meaning that a single word will activate many units in each representation layer, and a single unit in a layer is used in the representation of many different words. This is often implemented by considering the units

[5] A more complete model of lexical representation would include other levels as well, such as the syntactic category of the word, the pragmatic contexts which it is compatible with, and so on. Seidenberg & McClelland (1989) leave these levels out of the model because they are not as relevant as the phonological, orthographic, and semantic representations for the specific task of reading single words out of context.

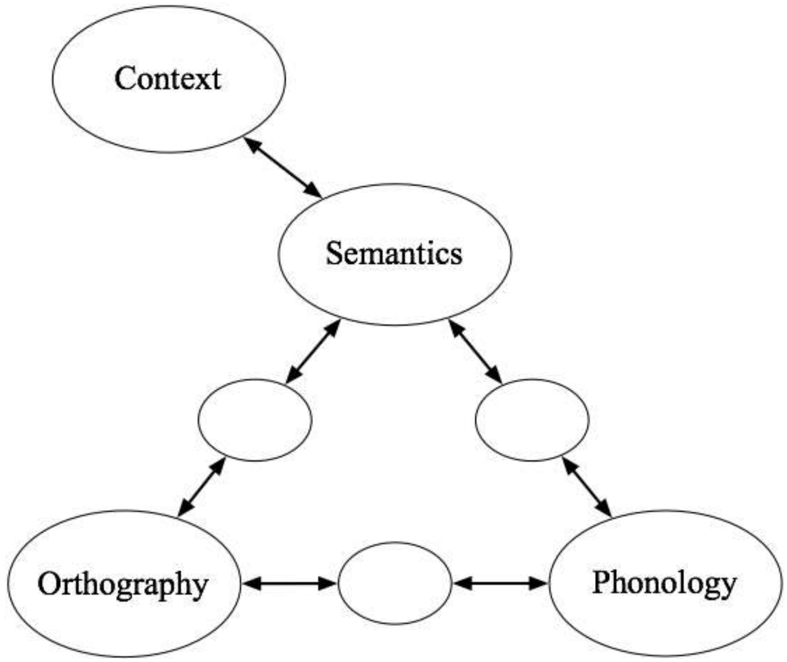


Figure 4.4: The Triangle Model, Seidenberg & McClelland's (1989) connectionist framework for lexical processing. Large ovals represent layers of representation units, while small ovals represent layers of hidden units mediating between representations. Arrows indicate bidirectional connections between layers.

in each layer to represent a set of binary features, with an activation level of $+1$ representing a $+$ value for the corresponding feature, and an activation of -1 representing a $-$ value. Second, activation is *interactive*, meaning that the activation levels of the units in one layer will both influence, and be influenced by, the activation at the other two layers. This is due to the bidirectional connections between the layers in the network.

The task of reading a single word in a TM network is accomplished by first setting the activations of the orthographic units to represent the spelling of the word. Then each unit in the network computes its activation level for a specified number of discrete time units. Because there are bidirectional connections between the different layers of the network, the representations formed at each level will influence the activations at the other two levels at each time step. Finally, the activation levels of the units in the phonology layer are read off as a phonological representation. In TM, lexical items are not represented on a distinct level of their own; rather, they exist as distributed mappings between orthographic, phonological, and semantic representations, and are encoded in the weighted connections between these three levels. This means that in a TM network, a word is read aloud by finding the best phonological representation which corresponds to the current orthographic representation, given the constraints on the mapping between orthography and phonology as encoded by the weights in the network. There are two routes by which this phonological representation for a word can be computed: the direct route from orthography to phonology, or the indirect route from orthography to semantics to phonology. Damage to one or the other of these routes in a trained TM network (as simulated by setting the values of randomly-chosen weights in either pathway to 0) can replicate the effects of different types of dyslexia seen in children (Plaut et al. 1996).

Plaut et al. (1996) provide a higher-level account of the origins of frequency and similarity effects in the TM framework. The basic phenomenon they are trying to account for is the interaction between frequency and consistency in word reading. Generally speaking, the more regular a spelling is, the faster and more accurately it will be read. However, this consistency effect decreases as the frequency of the word increases, so that there is very little difference in accuracy or latency between high-frequency words with regular or irregular spellings. Using a simplified version of one of their word-reading CNs,[6] they derive an equation showing how the effects of word frequency and similarity to previously-learned items interact during the processing of a given orthographic input

[6]The network they use to derive Equation 4.5 is a perceptron with a set of input orthographic units connected directly to a set of output phoneme units, with no intervening hidden units, and trained using a correlational (or "Hebbian"; Hebb 1949) learning rule, instead of an error-correcting rule like backpropagation. Plaut et al. point out that the behavior of an actual TM network would be more complex than this simplified network (but too complex to easily derive an equation describing its behavior). In particular, the hidden units and error correction during learning would allow a TM network, during the processing of an inconsistent lexical item, to overcome somewhat the negative effects of other, more consistent items in the lexicon that have a similar spelling. Equation 4.5 still serves as an approximation of a TM network's behavior, however.

(Plaut 2001; also see Goldrick 2007):

$$s_j^{[t]} = \sigma(F^{[t]} + \sum_f F^{[f]} O^{[ft]} - \sum_e F^{[e]} O^{[et]}). \tag{4.5}$$

Equation 4.5 says that the output s_j of the jth phoneme unit for a test pattern t is a sigmoidal function of the sum of three terms: the frequency $F^{[t]}$ that the test pattern occurred during the training of the network, the frequencies $F^{[f]}$ of the "friends" f of t during training (weighted by the degree of overlap $O^{[ft]}$ between the inputs for t and f), and the frequencies $F^{[e]}$ of the "enemies" e of t (again, weighted by the degree of overlap $O^{[et]}$). The friends of t are other lexical items that the network has been trained on which have the same output as t for unit j, while the enemies of t are lexical items which have the opposite output. Since this sum is input into a sigmoidal function, the friends and enemies of an input (which is where consistency effects derive from) will have less of an effect on high-frequency items than on low-frequency ones. This is because the output of a sigmoidal function does not linearly increase with increasing input, but instead grows closer to its maximum value (Figure 4.1). Assuming that accuracy and latency on an item depends on the strength of its output activations, this model would then predict that the regularity of spelling will have very little effect on high-frequency items, but more of an effect on lower-frequency items, with more consistent spellings being read faster and with greater accuracy than less consistent spellings.

While OT, HG, and TM are all examples of models which explain phonological regularities through a process of simultaneous constraint satisfaction, there are clearly some interesting differences between TM and generative models of phonology, as summarized in Table 4.1. One of the main differences between OT (and the Core-Periphery and Cophonology models) on the one hand, and HG and TM on the other hand, is the way in which constraints interact with each other. In OT constraints *strictly dominate* each other, such that no number of violations of a lower-ranked constraint can ever be worse than a single violation of a higher-ranked constraint. In HG, on the other hand, the

Table 4.1: Constraint-based models of phonology and the lexicon

Model	Constraint interaction	Lexicon topology	Frequency effects
Classic OT	strict domination	trivial	no
Core-Periphery	strict domination	partition	no
Cophonology	strict domination	partition	no
Harmonic Grammar	cumulative	trivial	no
Triangle Model	cumulative	metric	yes

constraints are given a numerical weight, and the badness of a form is calculated as the weighted sum of its constraint violations, which can allow (depending on the weights of the constraints in question) a form violating several lower-ranked constraints, or a single lower-ranked constraint multiple times, to be considered worse than a form violating a single higher-ranked constraint. In TM and other connectionist models, the constraints are encoded as the weights of the connections between units, and activation for a unit is computed as the weighted sum of incoming activation, which allows a similar kind of cumulative violation of constraints to occur (Dell 2000). Strict domination implies that there are never examples in any language of phonological patterns which show *cumulative markedness*, or "gang" effects, where a form that violates two or more lower-ranked constraints is considered worse than one which violates only one higher-ranked constraint. Prince & Smolensky (2006) claim that this is in fact the case for all known human languages. If true, this would be a surprising fact about the human language faculty, since in other cognitive domains, like vision, it does seem to be the case that the processing constraints involved interact cumulatively with each other (Dell 2000).

There are in fact some examples of phonological patterns which seem to show cumulative markedness effects. One example concerns the distribution of voiced geminates in Japanese loanwords. Kawahara (2006), citing corpus research by Nishimura (2003), shows that devoicing of voiced geminates may optionally occur in words where they appear with another voiced obstruent, such as [beddo] or [betto] 'bed', and [guddo] or

Table 4.2: Harmonic Grammar analysis of geminate devoicing in the presence of a voiced obstruent in Japanese loanwords (from Pater 2008)

doggɯ	1.5 IDENT-VOICE	1 *VCE-GEM	1 OCP-VOICE	
a. doggɯ		−1	−1	−2
b. ☞ dokkɯ	−1			−1.5

[gutto] 'good'. Devoicing cannot occur when the geminate occurs with a voiceless obstruent or with a sonorant, as in [kiddo], *[kitto] 'kid', and [webbɯ], *[weppɯ] 'web'. Kawahara analyzes this pattern as a conflict between two faithfulness constraints, one which preserves voicing in singletons and another which preserves voicing in geminates. However, Pater (2008) gives an alternative analysis using Harmonic Grammar, where the devoicing is a result of a cumulative constraint interaction between *VCE-GEM (no voiced geminates) and OCP-VOICE (no more than one voiced obstruent in a root).[7] Neither constraint is highly ranked enough to cause devoicing on its own, which allows Foreign items to have voiced geminates if there are no other voiced obstruents present in the word, as in [kiddo] 'kid'. However, when both constraints are violated, devoicing of the geminate can occur. Table 4.2 shows this gang effect between *VCE-GEM can OCP-VOICE for the word /doggɯ/ 'dog'. In this table, the top row indicates the numerical weights of each of the constraints, while the rightmost column shows the weighted sum of the constraint violations for each form. In this case, the form [dokkɯ] wins out over [doggɯ], despite it violating the higher-ranked constraint IDENT-VOICE, because the weight of IDENT-VOICE is less than the sum of the weights of *VCE-GEM and OCP-VOICE. Thus one violation of IDENT-VOICE is not enough to make the form [dokkɯ] worse than [doggɯ], which violates both *VCE-GEM and OCP-VOICE.[8]

[7] This constraint is also known as *Lyman's Law* in Japanese phonology, and governs the Yamato stratum (Chapter 1, section 1.2.1).

[8] Note that the weights given here for IDENT-VOICE, *VCE-GEM, and OCP-VOICE are somewhat arbitrary. Letting $w(\text{X})$ stand for the weight of constraint X, any set of weights satisfying the following three conditions will result in the geminate devoicing pattern: $w(\text{*VCE-GEM}) < w(\text{IDENT-VOICE})$,

But the existence or not of gang effects is only one way in which constraint-based theories of phonology differ. Another major difference between these theories concern their claims about the overall structure of the lexicon, and how different structures in the lexicon, such as lexical strata or smaller-scale lexical neighborhoods, may or may not interact with the phonological grammar. These differences in lexicon structure can be characterized as different kinds of *topologies*, where a topology on a set S is defined as a collection T of subsets of S (known as the *open sets* of S) obeying the following conditions (James 1999):

1. Both the empty set and S are open sets.
2. The intersection of any pair of open sets is also an open set.
3. The union of any number of open sets is also an open set.

Categorizing the claims about lexicon structure made by phonological theories as various kinds of topologies allows for them to be compared, in the sense that different topologies can entail relatively stronger or weaker notions of structure. A topology T' on a set S is said to be *finer* than a topology T on the same set (and T is *coarser* than T') if every open set of T is also an open set of T'. In addition, T' is *strictly finer* than T (and T *strictly coarser* than T') if $T \neq T'$; in other words, if there are at least some open sets in T' which are not also open sets in T. It turns out that the Core-Periphery and Cophonology models postulate a strictly finer topology on the lexicon than classic OT or HG, but a strictly coarser topology than TM. Since phonological properties of lexical items in the Core-Periphery and Cophonology models are tied to specific strata in the lexicon, this means that these models are able to account for different phonological processes applying to different groups of lexical items in a way that OT or HG cannot. However, the Core-Periphery and Cophonology models do not provide a rich enough

$w(\text{OCP-VOICE}) < w(\text{IDENT-VOICE})$, but $w(\text{*VCE-GEM}) + w(\text{OCP-VOICE}) > w(\text{IDENT-VOICE})$.

topological structure to be able to distinguish between lexical items in the same stratum, and so are not able to account for similarity effects governing the processing of individual lexical items. This additional similarity structure is what the topology induced by the TM model provides.

First let us consider the case of HG and classic OT. Both of these theories make no specific claims about how the lexicon is organized; these are theories of constraint interaction only. There are no constraints which are applied only to certain words or certain classes of words, and the same constraint ranking is used to evaluate all lexical items. The null hypothesis under these theories, then, is that the lexicon is an unstructured collection of lexical entries from which underlying representations are taken in order to be evaluated by Con (the constraint component of the grammar). This corresponds to the *trivial* topology, in which the only open sets are the empty set and the set *S* of all lexical items (James 1999). Since this trivially satisfies the conditions above (hence the name), this is the weakest possible notion of lexicon structure that can be postulated.

The Core-Periphery and Cophonology models are both examples of the *partition* topology, in which the set *S* of lexical items is partitioned into a collection of disjoint subsets which are then used to generate the open sets of the topology (Steen & Seebach 1995). This is strictly finer topology than the trivial topology used in OT and HG, since the only open sets in the trivial topology are the empty set and *S*, while in a partition topology there are additional subsets of *S* which are also identified as open sets. In the case of the Core-Periphery model, the open sets are the *lexical constraint domains*, the sets of lexical items governed by each indexed faithfulness constraint (with the empty set and *S* also included to satisfy the first condition above). Lexical strata are then defined as the set differences between the different constraint domains, so that for example the set of Sino-Japanese items is defined as all of the lexical items which are governed by FAITH_{SJ} but not FAITH_{Y} (Ito & Mester 1999). In the Core-Periphery model, each of the constraint domains is a strict subset of either another constraint domain or of the lexicon as a whole. This means that, for any pair of constraint domains A and B, where $\text{FAITH}_A \gg \text{FAITH}_B$ (and thus $A \supset B$), $A \cup B = A$ and $A \cap B = B$, and so the entire collection of constraint domains satisfies conditions (2) and (3) above. Therefore, the constraint domain structure induced over the lexicon by the ranking of indexed faithfulness constraints is an example of a partition topology.

The Cophonology model may initially seem to have a richer notion of lexicon structure than the Core-Periphery model, since it lacks the requirement that lexical strata are in a strict subset relationship with one another, but it too turns out to be an example of a partition topology. In a cophonology account, the lexicon is subdivided into a number of non-overlapping strata, and these strata form the lowest layer of a grammar lattice, with distinct constraint rankings being associated either directly with one of the strata, or with the set union of more than one stratum (as in Figure 3.7 in Chapter 3). This defines a partition topology in which the open sets are the subsets of the lexicon governed by the constraint rankings at each point of the grammar lattice (again, including the empty set and the set S of all lexical items). The union of any two of these open sets can always be found by following the lines of the grammar lattice upward to where they intersect, guaranteeing that the union of any number of open sets is itself an open set. The intersection of a pair of open sets can likewise be found by following the lines downward. This lattice structure thus shows that conditions (2) and (3) above are satisfied by the collection of partial ranking subsets.

By defining a partition topology over the lexicon, both the Core-Periphery and Cophonology models are able to identify subsets of the lexicon which have their own distinct phonological properties not shared by other lexical items (such as obligatory post-nasal voicing in the Yamato stratum, or the acceptability of /ti/ sequences in the Foreign stratum). In addition, a partition topology allows the relative closeness of a point to a subset of the lexicon to be defined, which is important for the Core-Periphery model, in which the nativization of a loanword is represented by the word moving closer to the core stratum of the lexicon. However, a partition topology alone is not a rich enough structure to define a notion of distance or similarity between two items. In other words, in either model, a lexical item A can be identified as belonging or not belonging to the same open set as another lexical item B, but there is no way to express the degree of similarity between A and B in terms of the topological structure alone. In addition, if A and B happen to be in the same open set, and there are no other open sets which contain A only and not B (or vice versa), then there is no way for them to be distinguished topologically, and thus no way for a phonological process to affect A without also affecting B.

Hence a partition topology turns out to still be a relatively weak notion of structure, meaning that there is no way in the Core-Periphery or Cophonology models to directly account for the phonological similarity effects that are seen over time in adaptation patterns like the adaptation of coronal obstruents before /i/, as discussed in Chapter 2. For this, it is necessary to make an even stronger assumption about how the lexicon is structured. Instead of partitioning the lexicon into a collection of disjoint subsets, a real-valued function $\sigma(A,B)$ can be defined over any pair of lexical items A and B, giving the degree to which a lexical item A is similar to B.[9] This forms the basis for a *metric* topology on the lexicon, where the open sets of S are generated from the *neighborhoods*

[9]Note that in an OT model, it is possible to define a similarity metric between any two lexical items based on the relative ranking of faithfulness constraints in the grammar and the number of violations of such constraints that would be required to change one of the items to the other. This would be analogous to the notion of *edit distance* used in information theory and computer science (Hamming 1950, Levenshtein 1966, Gusfield 1997, Navarro 2001). However, this faithfulness-based similarity relation is only used to help select between surface forms for a given UR; it is not used to compare distinct lexical items with each other.

of each item A in S, the set of items B such that $\sigma(A,B) > \varepsilon$ for a given ε. More formally, a metric space is defined using a real function $\delta(A,B)$ (the *metric* or *distance* function) which obeys the following three properties (Searcóid 2006):

1. Positivity: $\delta(A,B) \geq \delta(A,A) = 0$;
2. Symmetry: $\delta(A,B) = \delta(B,A)$;
3. Triangle inequality: $\delta(A,B) \leq \delta(A,C) + \delta(C,B)$.

Then the similarity $\sigma(A,B)$ between two items can be defined as an exponentially decreasing function of the distance between the items (Shepard 1987):

$$\sigma(A,B) = e^{-c\delta(A,B)}. \tag{4.6}$$

Connectionist models like TM which use distributed representations implicitly impose this kind of metric structure on their input representations. This is because similar items will have more input units in common that are activated than less similar items, and so the activations on the hidden layer for similar items will tend to be more similar than the activations for less similar items. This similarity-based activation is represented by the degree-of-overlap terms in Plaut et al.'s (1996) frequency-consistency relation in Equation 4.5. This property of TM networks is what allows them to be able to account for similarity effects in lexical processing that the four OT-based frameworks in Table 4.1 cannot account for.

In addition to the metric topology defined on the lexicon, which is a richer notion of structure than trivial or partition topologies, the TM framework and other connectionist models of lexical processing differ in one other important respect to the OT-based frameworks, namely, the role of lexical frequency in affecting the processing of a word. While word frequency plays no role in constraint ranking or evaluation in classic OT, in the TM framework frequency effects arise during the learning process. The more times a word is presented to a TM network during training, the more weight changes will be made to improve the accuracy of the network on the processing of that word. At the end of training, the frequency of the lexical items the network was trained on will be implicitly encoded in the values of the network's weights and will have an effect on processing. This is represented in Equation 4.5 by the F terms for the current test item and the friends and enemies of the test item. Thus a TM network can naturally account for frequency effects which cannot easily be explained using an OT-based model of phonology.

Note that certain kinds of frequency effects, namely the attested rates of different forms that are in free variation with one another, can be modeled in some versions of OT in which constraints have numerical weights whose values are gradually learned using a training algorithm, such as Stochastic OT (Boersma & Hayes 2001). The Cophonology model is also intended as an account of the frequencies of free variants, in that it predicts that these frequencies should correspond to the number of total rankings derived from the partial ranking of constraints from the grammar which can generate each form (Anttila & Cho 1998). However, these frameworks implicitly use either the trivial topology or a partition topology for the lexicon, and thus cannot also model effects on lexical processing deriving from the frequencies of particular items or environments in the lexicon itself, such as the likelihood of a non-native adaptation depending on the type frequency of the phonological environment it occurs in. The TM framework is better able to account for these kinds of effects as well, as evidenced by Plaut et al.'s (1996) results, and thus is a more promising model for accounting for these types of frequency effects on loanword adaptation patterns.

To summarize, the TM framework, like OT-based frameworks, is a constraint-based model of phonological processing. However, it differs in three important respects from classic OT. First, constraints interact cumulatively rather than strictly dominating each other, allowing for gang effects to occur. Second, a TM network imposes a similarity metric on its input representations, which means the processing of a lexical item can be influenced by previously-learned items that are sufficiently similar to it. Finally, the frequency of an item during the training of a network also plays a role in its processing, which provides an explanation for the frequency-consistency interaction. In the remainder of this chapter, I will investigate the usefulness of this framework in explaining some of the loanword adaptation patterns from Chapter 2, focusing in particular on the adaptation of voiced geminates.

4.3 A connectionist framework for lexical representation in Japanese

In order to test the viability of the TM framework for explaining frequency and similarity effects on loanword adaptation patterns, in the remainder of this chapter I develop two connectionist models of lexical representation in Japanese, based on the implementations of TM presented in Plaut et al. (1996) and Harm & Seidenberg (1999). The questions I seek to answer are the following:

1. What is the nature of the similarity structure of the Japanese lexicon, and how can this structure be exploited to learn stratal classifications?
2. How can loanword adaptation be simulated so that frequency and phonological similarity effects arise naturally from the mechanisms of the model itself?

Table 4.3: The major class features [son] and [cons]

	son	cons	voi	nas
Voiceless obstruents	−	+	−	−
Voiced obstruents	−	+	+	−
Nasals	+	+	+	+
Approximants	+	+	+	−
Vowels and glides	+	−	+	−

In the first model, the *stratum classification network* (section 4.4), I explore how the large-scale organization of the lexicon in terms of lexical strata can be represented in a CN. In the second model, the *loanword adaptation network* (section 4.5), I train several networks on a phonological repetition task, and then use these networks to look at the influence of lexical type frequency on the adaptation of voiced geminates in loanwords. In the remainder of this section, I will present the feature system and stratal classification algorithm used in generating the training data for both networks.

4.3.1 Feature representation

For both of the models described in the following two sections, a feature system representing the phonological contrasts in Japanese (Figure 4.5) is used to encode each mora of the training and test words as a numerical vector suitable for presentation to the network. The system of features I am using here is based on the feature geometry presented in Newman (1997), which is similar to that of Sagey (1990), except that it uses the account of palatalization developed in Lahiri & Evers (1991). The features [son] (sonorant), [cons] (consonantal), [voi] (voice), and [nas] (nasal) have similar definitions as in Chomsky & Halle (1968) and following work, as shown in Table 4.3. There are two separate [cont] (continuancy) features, [cont_1] and [cont_2]. These represent the continuancy feature at the beginning and end of the segment, respectively, and are used to

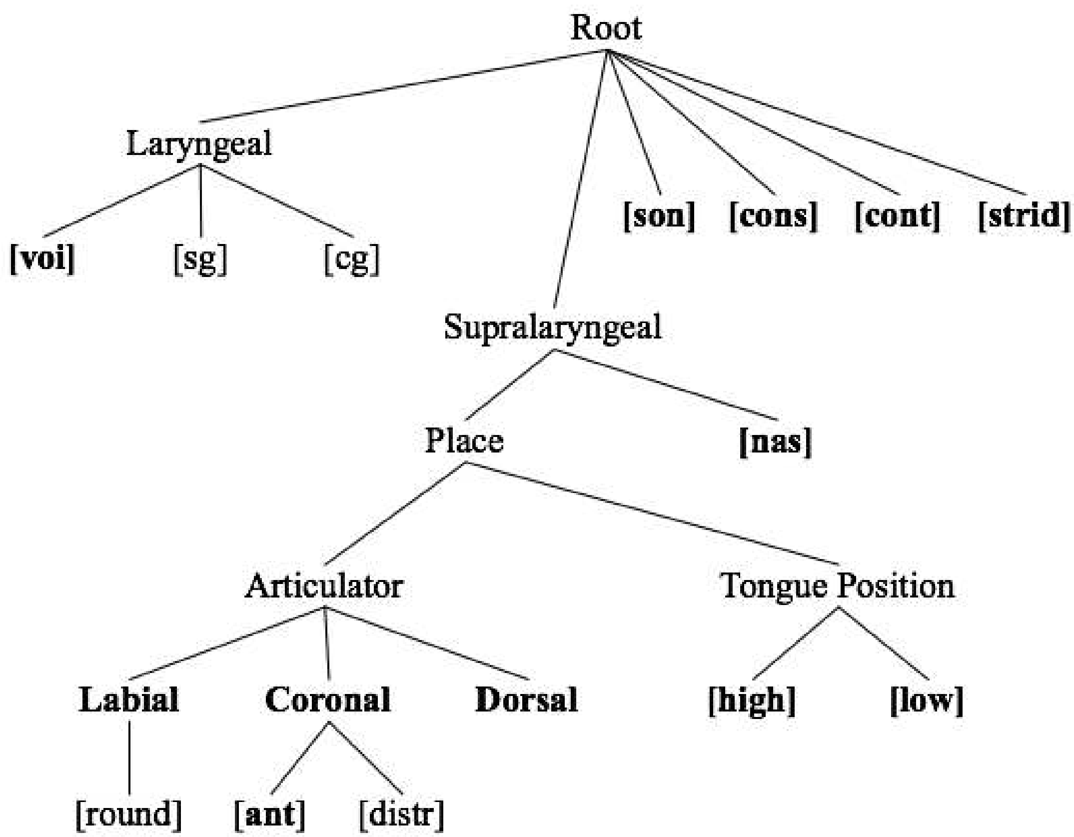


Figure 4.5: Features used in network simulations, based on the feature geometry presented in Newman (1997). Features used in network training are shown in bold.

Table 4.4: The features [cont_1] and [cont_2]

	cont_1	cont_2
Stops	−	−
Affricates	−	+
Fricatives	+	+

Table 4.5: Articulator features, [ant], and [high]

	lab	cor	dors	ant	high
Labials	+	−	−	−	−
Dentals	−	+	−	+	−
Palatals	−	+	−	−	+
Velars	−	−	+	−	−
Front vowels	−	+	−	−	±
Back vowels	−	−	+	−	±

distinguish between stops, affricates, and fricatives (Table 4.4). This resembles Sagey's (1990) analysis of affricates as complex segments with multiple values for the [cont] feature. The articulator features [lab], [cor], and [dors] represent whether the lips, tongue tip/blade, or tongue dorsum, respectively, are used in articulating the segment, and are used to distinguish between places of articulation (Table 4.5). However, contra Sagey (1990), these features are binary, not privative,[10] for the sake of keeping the input representation for the network relatively simple.

Following Lahiri & Evers (1991), the features [cor] and [dors] are also used to distinguish between front and back vowels. Front vowels are [+cor, −dors], while back vowels are [−cor, +dors].[11] In the Lahiri & Evers (1991) model, the features [high] and [low] are not under the Dorsal articulator node, as in Sagey (1990), but instead are placed under a new node called Tongue Position, which is present for all segments, not just dorsal segments. This allows palatalized segments to be consistently marked [+high],

[10]Privative features do not take + or − values; instead they are either present or absent in the segment.

[11]Note that Lahiri & Evers (1991) use privative articulator features, not the binary features I am using here. Thus they represent the difference between front and back vowels in terms of whether there is a Coronal or Dorsal node linked to the Articulator node.

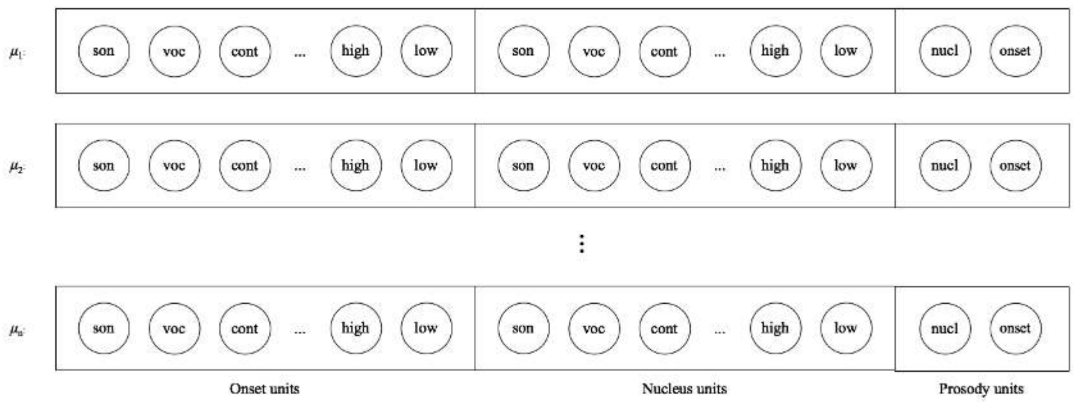


Figure 4.6: Slot-based representation for Japanese phonology

which allows for an elegant account of palatalization before high vowels in terms of spreading of the [high] feature from the vowel to the preceding consonant (Lahiri & Evers 1991).

All of the features are binary and are coded numerically using the values $+1$ and -1 to represent $+$ and $-$ values for each feature. These feature values for each of the segments in the word[12] are then concatenated into a single vector using a slot-based representation as shown in Figure 4.6. Table 4.6 gives the feature values for Japanese segments, while Table 4.7 provides some examples of Japanese words encoded using this system. This representation is essentially equivalent to the sequences of unstructured feature bundles used in Chomsky & Halle (1968). The dependency relations between features discovered in later research on feature geometry (Sagey 1990, McCarthy 1988, Clements & Hume 1995) are not directly represented. This is because it is a non-trivial task to represent these hierarchical structures as a numerical vector. The dominance relationships between the individual features would have to be encoded as values in the vector somehow, for example as a tensor product representation (Smolensky 1990,

[12]The words are encoded using a narrow phonetic transcription. Thus geminates are represented as two segments ([tt], [kk], and so on), while the placeless nasal /ɴ/ is represented as either [m], [n], [ɲ], [ŋ], [ɴ], or a nasalized vowel, depending on its position in the word, and the following segment, if any (Akamatsu 1997).

2006a,c) or using a convolution operation (Plate 1994; see also Steedman 2001). For the sake of simplicity, I have instead chosen to represent segments as feature vectors, and allow each network to learn the dependencies between features as necessary for the particular task it is being trained on.

There are two additional units for each mora slot which are used only in the loanword adaptation network in section 4.5. These are labelled [nucl] and [onset] in Table 4.7. These two units encode the presence of a nucleus and an onset, respectively, in the mora slot they are associated with. If the value of [nucl] is −, then the mora slot is empty, whereas if it is +, then there is a mora present in the mora slot. Likewise, if the value of [onset] is −, then the mora is an onsetless vowel, the placeless nasal /N/, or the first segment /Q/ of a geminate consonant, whereas if [onset] is +, then the mora is CV. These two features are actually redundant, since their values can be predicted from the values of the other features in the mora slot. They have been added for the adaptation network to specify the prosodic structure that the loanword adaptation should have (specifically, whether the network should try to produce a geminate or not).

4.3.2 Stratum classification

In order to generate the training and test data for the models in the next two sections, it was necessary to classify the entries in `JMDICT` (Electronic Dictionary Research and Development Group 2003) by lexical stratum. This was done on the basis of each entry's orthography, according to the following algorithm:

1. Let w = entry from `JMDICT`
2. If w is an adverb or verb, and is formed from a bimoraic root which is either reduplicated or has a suffix −ɾi, −(ɾi)to, −tsɯkɯ, −mekɯ, −nakɯ, −(to)sɯɾɯ,

Table 4.6: Feature encoding for Japanese segments. The list of segments is from Akamatsu (1997). [c] and [ɟ] occur only as the first part of a geminate /ʧ/ (=[cɕ]) and /ʤ/ (=[ɟʑ]). [ŋ] and [ŋʲ] occur as allophones of /ɴ/ before velar stops, while [tʲ] and [dʲ] occur only in loanwords.

	Manner							Place					
Segment	son	cons	$cont_1$	$cont_2$	strid	nas	voi	lab	cor	dors	ant	high	low
Ø	−	−	−	−	−	−	−	−	−	−	−	−	−
p	−	+	−	−	−	−	−	+	−	−	−	−	−
pʲ	−	+	−	−	−	−	−	+	−	−	−	+	−
t	−	+	−	−	−	−	−	−	+	−	+	−	−
tʲ	−	+	−	−	−	−	−	−	+	−	+	+	−
t͡s	−	+	−	+	+	−	−	−	+	−	+	−	−
c	−	+	−	−	−	−	−	−	+	−	−	+	−
cɕ	−	+	−	+	+	−	−	−	+	−	−	+	−
k	−	+	−	−	−	−	−	−	−	+	−	−	−
kʲ	−	+	−	−	−	−	−	−	−	+	−	+	−
b	−	+	−	−	−	−	+	+	−	−	−	−	−
bʲ	−	+	−	−	−	−	+	+	−	−	−	+	−
d	−	+	−	−	−	−	+	−	+	−	+	−	−
dʲ	−	+	−	−	−	−	+	−	+	−	+	+	−
ɟ	−	+	−	−	−	−	+	−	+	−	−	+	−
ɟʑ	−	+	−	+	+	−	+	−	+	−	−	+	−
ɡ	−	+	−	−	−	−	+	−	−	+	−	−	−
ɡʲ	−	+	−	−	−	−	+	−	−	+	−	+	−
ɸ	−	+	+	+	−	−	−	+	−	−	−	−	−
s	−	+	+	+	+	−	−	−	+	−	+	−	−
z	−	+	+	+	+	−	+	−	+	−	+	−	−
ɕ	−	+	+	+	+	−	−	−	+	−	−	+	−
ç	−	+	+	+	−	−	−	−	+	−	−	+	−
h	−	+	+	+	−	−	−	−	−	−	−	−	+
m	+	+	−	−	−	+	+	+	−	−	−	−	−
mʲ	+	+	−	−	−	+	+	+	−	−	−	+	−
n	+	+	−	−	−	+	+	−	+	−	+	−	−
nʲ	+	+	−	−	−	+	+	−	+	−	+	+	−
ɲ	+	+	−	−	−	+	+	−	+	−	−	+	−
ŋ	+	+	−	−	−	+	+	−	−	+	−	−	−
ŋʲ	+	+	−	−	−	+	+	−	−	+	−	+	−
ɴ	+	+	−	−	−	+	+	−	−	+	−	−	+
ɾ	+	+	+	+	−	−	+	−	+	−	+	−	−
ɾʲ	+	+	+	+	−	−	+	−	+	−	+	+	−
a	+	−	+	+	−	−	+	−	−	+	−	−	+
i, j	+	−	+	+	−	−	+	−	+	−	−	+	−
ɯ, ɰ	+	−	+	+	−	−	+	−	−	+	−	+	−
e	+	−	+	+	−	−	+	−	+	−	−	−	−
o	+	−	+	+	−	−	+	−	−	+	−	−	−

Table 4.7: Example vector representations for training items. In mora column, O=onset, N=nucleus.

Word		Mora	Manner							Place						Prosody	
			son	cons	$cont_1$	$cont_2$	strid	nas	voi	lab	cor	dors	ant	high	low	nucl	onset
/aka/ 'red'	μ_1	O: ø	−1	−1	−1	−1	−1	−1	−1	−1	−1	1	−1	−1	−1		
		N: ø	−1	−1	−1	−1	−1	−1	−1	−1	−1	1	−1	−1	−1	−1	−1
	μ_2	O: ø	−1	−1	−1	−1	−1	−1	−1	−1	−1	1	−1	−1	−1		
		N: ø	−1	−1	−1	−1	−1	−1	−1	−1	−1	1	−1	−1	−1	−1	−1
	μ_3	O: ø	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1		
		N: a	1	−1	1	1	−1	−1	1	−1	−1	1	−1	−1	1	1	−1
	μ_4	O: k	−1	1	−1	−1	−1	−1	−1	−1	−1	1	−1	−1	−1		
		N: a	1	−1	1	1	−1	−1	1	−1	−1	1	−1	−1	1	1	1
/ʧotto/ 'a little'	μ_1	O: ø	−1	−1	−1	−1	−1	−1	−1	−1	−1	1	−1	−1	−1		
		N: ø	−1	−1	−1	−1	−1	−1	−1	−1	−1	1	−1	−1	−1	−1	−1
	μ_2	O: cɕ	−1	1	−1	1	1	−1	−1	−1	1	−1	−1	1	−1		
		N: o	1	−1	1	1	−1	−1	1	−1	−1	1	−1	−1	−1	1	1
	μ_3	O: ø	−1	−1	−1	−1	−1	−1	−1	−1	−1	1	−1	−1	−1		
		N: t	−1	1	−1	−1	−1	−1	−1	−1	1	−1	1	−1	−1	1	−1
	μ_4	O: t	−1	1	−1	−1	−1	−1	−1	−1	1	−1	1	−1	−1		
		N: o	1	−1	1	1	−1	−1	1	−1	−1	1	−1	−1	−1	1	1
/beɴkʲoː/ 'study'	μ_1	O: b	−1	1	−1	−1	−1	−1	1	1	−1	−1	−1	−1	−1		
		N: e	1	−1	1	1	−1	−1	1	−1	1	−1	−1	−1	−1	1	1
	μ_2	O: ø	−1	−1	−1	−1	−1	−1	−1	−1	−1	1	−1	−1	−1		
		N: ŋʲ	1	1	−1	−1	−1	1	1	−1	−1	1	−1	−1	−1	1	−1
	μ_3	O: kʲ	−1	1	−1	−1	−1	−1	−1	−1	−1	1	−1	1	−1		
		N: o	1	−1	1	1	−1	−1	1	−1	−1	1	−1	−1	−1	1	1
	μ_4	O: ø	−1	−1	−1	−1	−1	−1	−1	−1	−1	1	−1	−1	−1		
		N: o	1	−1	1	1	−1	−1	1	−1	−1	1	−1	−1	−1	1	−1

or −(to)kɯɾɯ, then assign *w* to Mimetic stratum

3. If *w* is written only with *katakana*, then assign to Foreign stratum
4. If *w* is written only with *hiragana*, then assign to Yamato stratum
5. If *w* is written in a mixture of *katakana* and *hiragana*, or *katakana* and *kanji*, then assign to Hybrid-F stratum
6. If *w* has not been assigned to a stratum yet, generate all possible *on-yomi* and *kun-yomi* for each of the *kanji* characters in *w* (at this point we know *w* is written either only in *kanji*, or in a mixture of *kanji* and *kana*)
7. If the actual reading for *w* matches one of the generated *on-yomi*, assign to Sino-Japanese
8. If the actual reading for *w* matches one of the generated *kun-yomi*, assign to Yamato
9. Otherwise, assign *w* to Hybrid-YS stratum (because at this point we know there are both *on-yomi* and *kun-yomi kanji* in *w*)

The Yamato, Sino-Japanese, and Foreign strata are easily distinguished by the scripts normally used for each: *katakana* for Foreign, *kanji* for Sino-Japanese, and *hiragana* and *kanji* for Yamato.[13] Moreover, *kanji* characters generally have two sets of readings, the Sino-Japanese reading (*on-yomi*), and the Yamato reading (*kun-yomi*). For example, the character 山 is used both for the Sino-Japanese root /saɴ/ and the Yamato word /jama/, both meaning 'mountain'. Thus a word written in *kanji* can be classified as

[13]This is what makes the study of lexical strata in Japanese so much more tractable than in other languages: the complexity of the writing system and the fact that each stratum has its own distinct set of scripts associated with it. It would be much more difficult to write a program to classify all of the entries in an English dictionary as either native, Latinate, or recent loanwords (although still possible, given how English generally preserves the original orthography in Latin/Greek and recent loanwords); while doing this for a language with a shallow orthography, like Spanish or Finnish, would necessitate classifying words in terms of their phonological and phonotactic characteristics, which would be considerably more error-prone for stratum identification.

Sino-Japanese or Yamato based on whether the *on-yomi* or *kun-yomi* readings of the characters are used. There are also many words in JMDICT composed of morphemes from different strata; these are classified as either Hybrid-F or Hybrid-YS. A word written in both *katakana* and *kanji*, or *katakana* and *hiragana*, is classified as Hybrid-F. A word which uses *kun-yomi* for some of its *kanji*, and *on-yomi* for the other *kanji*, is classified as Hybrid-YS. For entries with more than one possible spelling, the most frequent spelling was used to classify by stratum, as determined from a frequency list (Kamermans 2008) made available through the JMDICT project. There are some high-frequency words, like 為る /sɯɾɯ/ 'to do' and 成る /naɾɯ/ 'to become', which are usually spelled using *kana* only (する and なる, respectively). These words are specially marked as such in JMDICT and were classified as Yamato by rule 4 above.

Using these orthographic criteria amounts to using etymology to classify each word in the dictionary. This largely corresponds to the synchronic classification, with regards to processes like *rendaku*, although as I noted before in Chapter 1 there are a few exceptional words like 本 /hoɴ/ 'book' which were originally borrowings but now act like they are a part of the Yamato stratum. Words like these will be (perhaps incorrectly, depending on one's point of view) classified as Sino-Japanese or Foreign by the stratum-classification algorithm.

Unfortunately, there is no way to distinguish the Mimetic items using orthographic criteria, because in JMDICT they are written using either *hiragana* or *katakana*. However, Mimetic words are usually formed from bimoraic roots, which are either reduplicated or occur with a suffix like /−ɾi/, /−tsɯkɯ/, or /−tosɯɾɯ/ (Hamano 1998), and so all words with these properties are classified as Mimetic by the algorithm.

Table 4.8 summarizes the results of the stratum-assignment algorithm applied to the entries in JMDICT. The proportions of Yamato, Sino-Japanese, and Foreign found in

Table 4.8: Lexical strata derived from JMDICT

Yamato	26,200	(19%)
Sino-Japanese	61,566	(45%)
Hybrid-YS	13,409	(10%)
Mimetic	815	(1%)
Foreign	21,824	(16%)
Hybrid-F	944	(1%)
Unknown	12,711	(9%)
Total	137,469	

JMDICT are similar to the figures cited in Shibatani (1999: 142–3) for the number of distinct lexical items in magazines and newspapers published in the 1950's and 60's, except that there are fewer Yamato words in JMDICT (compared to about 30–40% in the texts Shibatani cites), and more Foreign words (compared to about 10% in Shibatani). Presumably this is because the dictionary lists many foreign borrowings which are not commonly found in real texts, and so this skews the ratios towards these types of items. Also, the number of Mimetic items found is extremely small, because such terms are usually not listed in Japanese dictionaries (Hamano 1998) and because of the difficulty identifying them as noted above. Since there was not a large enough sample of Mimetic words, in the two networks that follow, only words classified as either Yamato, Sino-Japanese, or Foreign are used in the training and test sets. Also, to make the training simpler, Hybrid-F and Hybrid-YS words were not used either.

4.4 Stratum classification network

The first network was designed to investigate questions raised by Ota (2004) (and earlier by Bloch 1950 and Rice 1997) about the learnability of lexical stratification in Japanese. Ota notes that in the process of language acquisition, it seems a child should not be able to learn the correct stratal classifications for the words it is exposed to in the absence of evidence from alternations. This is because there are many words, such as 天馬 /teɴba/ 'flying horse; Pegasus' and コンボ /koɴbo/ 'combo', which do not violate the static distributional constraints on Yamato roots and yet are not themselves members of the Yamato stratum (the two examples given are Sino-Japanese and Foreign, respectively). Ota points out that:

> [w]hile the analyst, equipped with etymological knowledge, may be able to assign all morphological items to the different sublexica... it is doubtful that the same stratified lexicon can be reconstructed in a bottom-up fashion if the membership of some items can be determined only on the basis of surface distribution pattern.... [E]ven if we can justify the existence of phonological sublexica, we cannot determine the classhood of all lexica based solely on distributional evidence. This problem has severe implications for the acquisition of nonuniform phonology.... (Ota 2004: 23)

Ota concludes that the correct stratal classification of a word cannot be reliably made using distributional evidence only. Alternations (such as past tense voicing) must also be taken into account in many cases to determine which stratum a word belongs to. But this raises the question, just how far can we get using only distributional evidence? In other words, how well can a word be classified by stratum purely on the basis of its phonological form? It may be that distributional evidence is enough to learn many or even most of the correct stratum classifications. One reason to suspect this is the case is the distribution of syllable types in the different strata. Table 4.9 shows the distribution of weak (monomoraic) and strong (bimoraic) syllables in all of the four-mora words in the Yamato, Sino-Japanese, and Foreign strata listed in `JMDICT`. While Foreign words are evenly distributed among the five possible syllable parses for four-mora words, Yamato items show an overwhelming tendency to be composed of weak syllables only, while

Table 4.9: Distribution of syllable types for four-mora words in the Yamato, Sino-Japanese, and Foreign strata

Syllable parse	Yamato		Sino-Japanese		Foreign	
$(\mu)(\mu)(\mu)(\mu)$	6,794	(81%)	1,142	(6%)	801	(24%)
$(\mu\mu)(\mu)(\mu)$	556	(7%)	4,235	(22%)	532	(16%)
$(\mu)(\mu\mu)(\mu)$	505	(6%)	643	(3%)	791	(24%)
$(\mu)(\mu)(\mu\mu)$	396	(5%)	2,405	(12%)	652	(19%)
$(\mu\mu)(\mu\mu)$	87	(1%)	11,032	(57%)	579	(17%)

Sino-Japanese items show the opposite tendency to have at least one strong syllable in the word. Ultimately the reason Sino-Japanese items often have strong syllables is because they are composed of roots which were borrowed from Middle Chinese closed monosyllables. A Sino-Japanese root can be either a single syllable with a rime of the form /V/, /Vː/, /VN/, /VQ/, or /Vi/, or two weak syllables, the second of which must be one of /ki/, /kɯ/, /ti/, or /tɯ/[14] (Martin 1952, Tateishi 1990).[15] These different syllable types reflect different ways to repair coda consonants which were not allowed in Classical Japanese. Yamato items, on the other hand, are descended from Old Japanese, which only allowed CV syllables (Martin 1987, Frellesvig 1995),[16] thus explaining the preponderance of weak syllables in modern-day Yamato words.

Of course these differences in syllable parses are statistical generalizations only, since examples of four-mora words with any possible combination of syllable types can be found in all three strata. But these tendencies do suggest that there is some distributional evidence available to the Japanese learner which can be exploited to learn stratal classifications. In this section I will explore the question of how lexical strata

[14]The syllables /ti/ and /tɯ/ in Sino-Japanese roots surface as [ʧi] and [tsɯ], respectively.

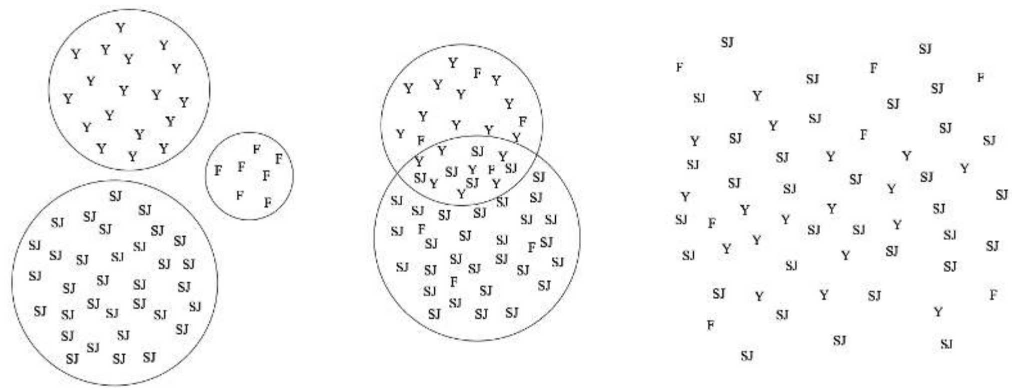
[15]Martin (1952) actually lists all of the possible Sino-Japanese morpheme forms, rather than stating the distribution as a set of rules. An examination of his list shows that there may be even more specific constraints; for example, bisyllabic morphemes beginning with /b/, /ʤ/, or /n/ apparently only occur with /kɯ/ or /tɯ/ in the second syllable, never /ki/ or /ti/. This is probably a historical accident, though, since /b/, /ʤ/, and /n/ do not form a natural class.

[16]Although there may have been some Sino-Japanese items in Old Japanese with syllables of the form C(G)V(G), where G is a glide /j,w/ (Miyake 2003).

can be learned on the basis of distributional evidence by constructing a feedforward CN classifier and evaluating its performance on a random sample of lexical items chosen from `JMDICT`.

4.4.1 The similarity structure of the Japanese lexicon

I will begin to answer the question of how easily the lexical stratum of a word can be predicted from its phonological form by reframing this question as a slightly different problem. Suppose each item in the Japanese lexicon is encoded mora-by-mora according to the feature system presented in section 4.3.1. Then each word of length n moras will be represented as a vector of length $26n$ (since each mora is represented by 26 feature values, 13 for the nucleus and 13 for the optional onset). Alternatively, each word is represented by a single point in a $26n$-dimensional vector space. The question now is, for all words of length n, can distinct regions be identified in this vector space corresponding to each lexical stratum, such that the majority of Yamato (or Sino-Japanese or Foreign) items are found inside this region, while most non-Yamato (or Sino-Japanese or Foreign) items are found outside. The degree to which distinct regions corresponding to lexical strata can be identified should then correspond to the learnability of stratal classifications. If the points corresponding to Yamato, Sino-Japanese, and Foreign items are easily separable (as shown schematically in Figure 4.7a) then it should be easy to decide, solely on the basis of a word's phonological form, which stratum it is a member of. Conversely, if there are no identifiable regions in which only Yamato, Sino-Japanese, or Foreign items occur (Figure 4.7c), then it will be impossible to guess which stratum a particular word is a member of without at least also considering evidence from morphological alternations. Other kinds of lexicon organization patterns are also possible that are intermediate between these two extremes. For example, Figure 4.7(b) depicts a



(a) Completely distinct strata

(b) Partially distinct strata

(c) No distinguishable strata

Figure 4.7: Schematic representations of possible lexicon organization patterns

situation in which Yamato and Sino-Japanese items are relatively (though not perfectly) distinguishable, while Foreign items are interspersed throughout the space occupied by the other two strata, making it perhaps more difficult, on the basis of phonological structure alone, to identify a given word as being Foreign.[17]

As a way of visualizing whether the structure of the Japanese lexicon is more like (a), (b), or (c) in Figure 4.7, I used an exploratory data-analysis technique known as *multidimensional scaling* (MDS; Kruskal & Wish 1978, Cox & Cox 2001), which takes a matrix of distances between a set of points in a high-dimensional space, and produces a projection of the points into a lower-dimensional space (usually two- or three-dimensional) in which the relative distances between the original set of points is preserved as much as possible. This allows the proximity structure of the original high-dimensional data to be more easily visualized. In this case, I took a random sample of 5000 Yamato, Sino-Japanese, and Foreign words from `JMDICT` that were from one to four moras in length and occurred at least 50 times in a frequency database of words occurring in Japanese novels (Kamermans 2008). Each word was encoded mora-by-mora according

[17]Of course, this only applies to loanwords like /koNbo/ 'combo' which do not contain any non-native phonology. If a loanword does happen to contain a non-native pattern such as /ti/, then it will be easily identified as a member of the Foreign stratum.

to the feature system presented in section 4.3.1. For four-mora words, this created a 104-element[18] vector representing the feature values of the segments in each of the four moras, aligned so that the features for the nuclei of each mora occur in the same place in the vector for every word. For words that were less than four moras in length, empty moras /Ø/ (which are encoded in the vector by setting all of the feature values to -1) were appended on the left until the word was four moras long. Then the distance matrix for all of the words in the sample was generated by taking the Euclidean distance $\delta(A,B)$ between the vector representations for each pair of words A and B in the sample:

$$\delta(A,B) = \sqrt{\sum_{i=1}^{n}(A_i - B_i)^2}, \tag{4.7}$$

where A_i denotes the ith element of the vector representation for word A. I then performed MDS on this distance matrix using the `isoMDS` function from the `MASS` package (Venables & Ripley 2002) for the `R` statistical programming environment (R Development Core Team 2006). The results are shown in Figure 4.8. In this case, I had to use a four-dimensional MDS solution, since I found that the stress values (the discrepancy between the original distances, and the distances of the projected points) were too high (> 0.15) for two- and three-dimensional solutions. The four rows and columns of the matrix of scatterplots in Figure 4.8 correspond in order to the four dimensions of the MDS projection, and each cell of the matrix contains a scatterplot using two of the four dimensions (the dimensions corresponding to the row and column that the cell is in).

It is clear from Figure 4.8 that the Yamato items (represented by the red points in the scatterplots) in the sample tend to occur in one of two main clusters, while Sino-Japanese items (represented by the blue points) are grouped into four or five smaller clusters. This clustering structure is more evident in Figure 4.9, showing a scatterplot of MDS axes 1 and 2, and Figure 4.10, a scatterplot of axes 1 and 3. (These two figures correspond to

[18]Because there are thirteen features for each segment, each mora is represented by 26 features; $26 \times 4 = 104$.

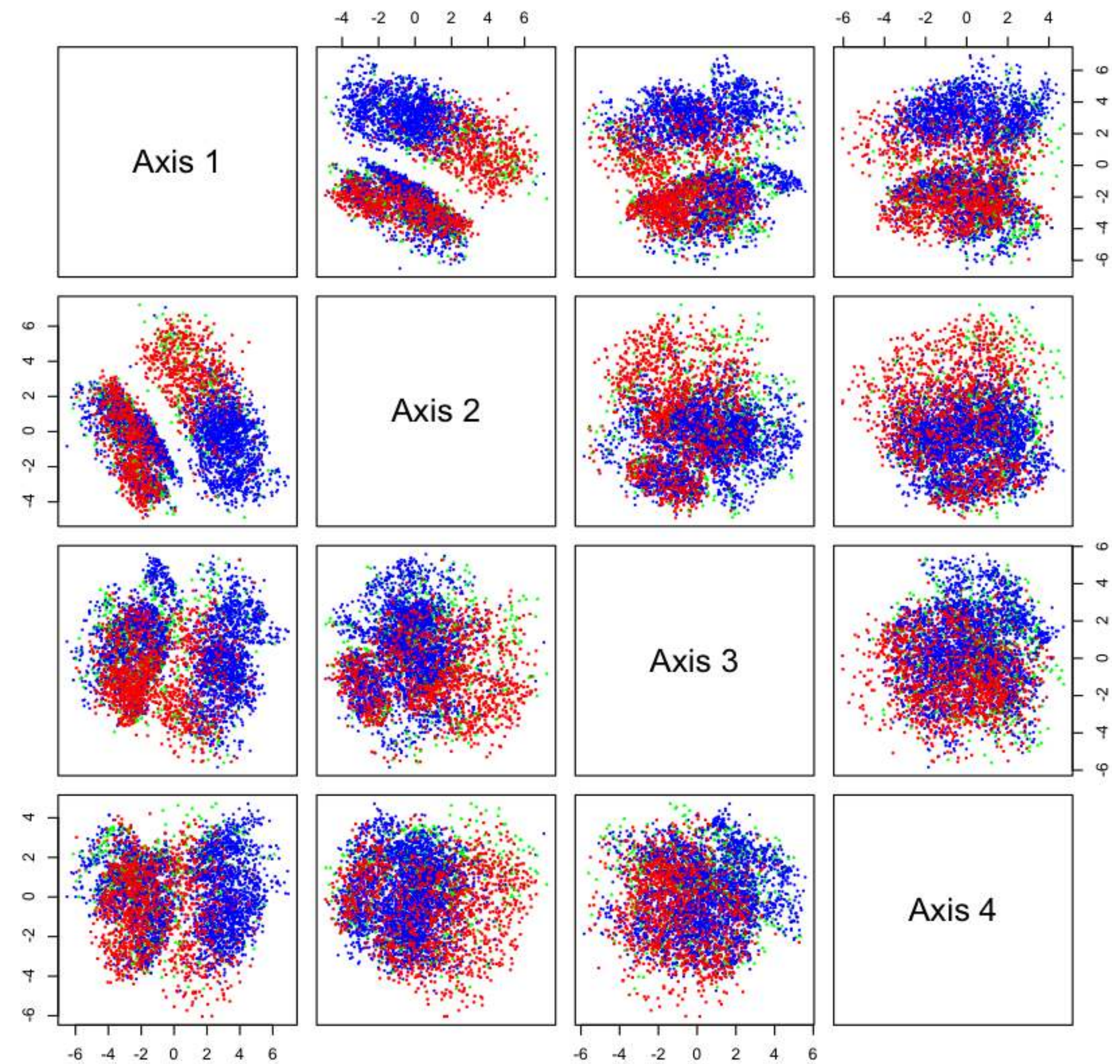


Figure 4.8: MDS analysis of the feature vector representations for a random sample of 5000 lexical items from JMDICT ($k = 4, \text{stress} = 0.1492$). Each cell in the grid contains a scatterplot using two of the four MDS axes. Red points are Yamato items, blue Sino-Japanese, and green Foreign.

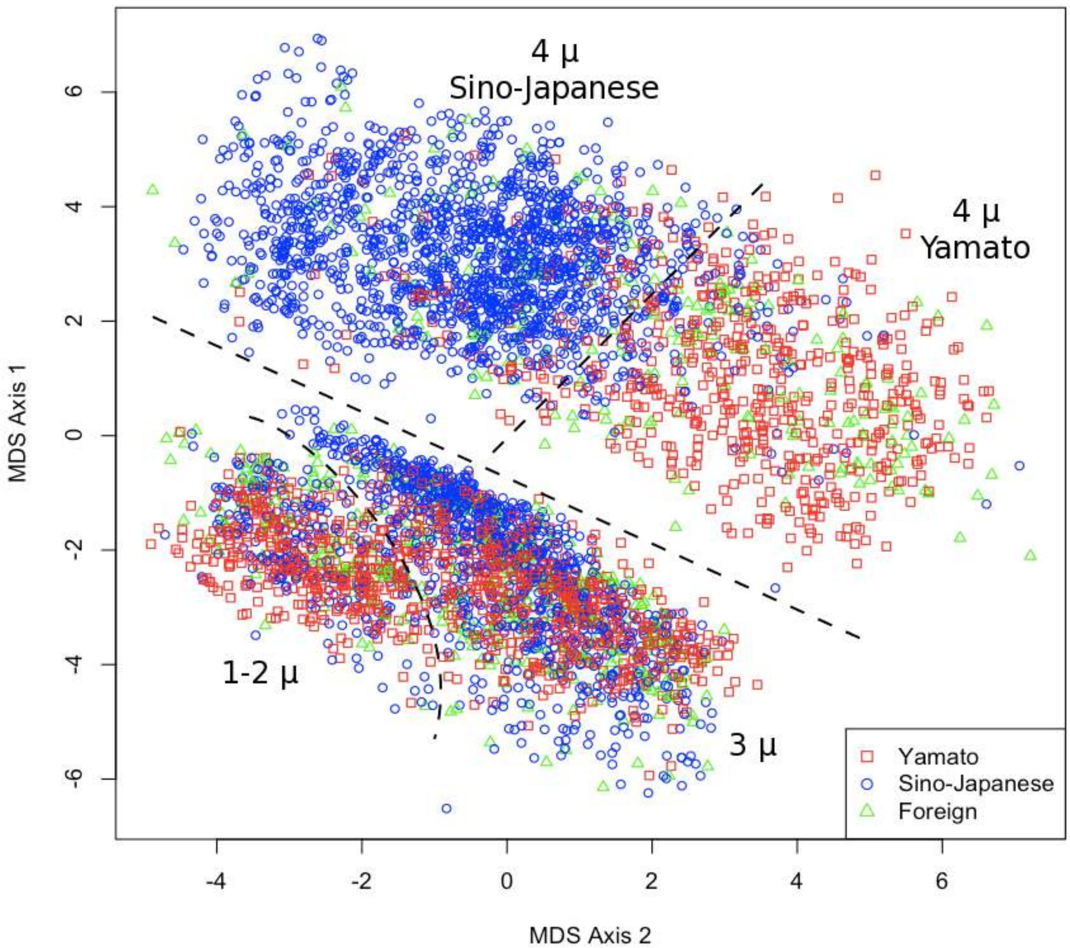


Figure 4.9: Scatterplot of MDS axes 1 and 2 from Figure 4.8, labeled with clusters corresponding to four-mora Yamato/Foreign words, four-mora Sino-Japanese words, and words with less than four moras.

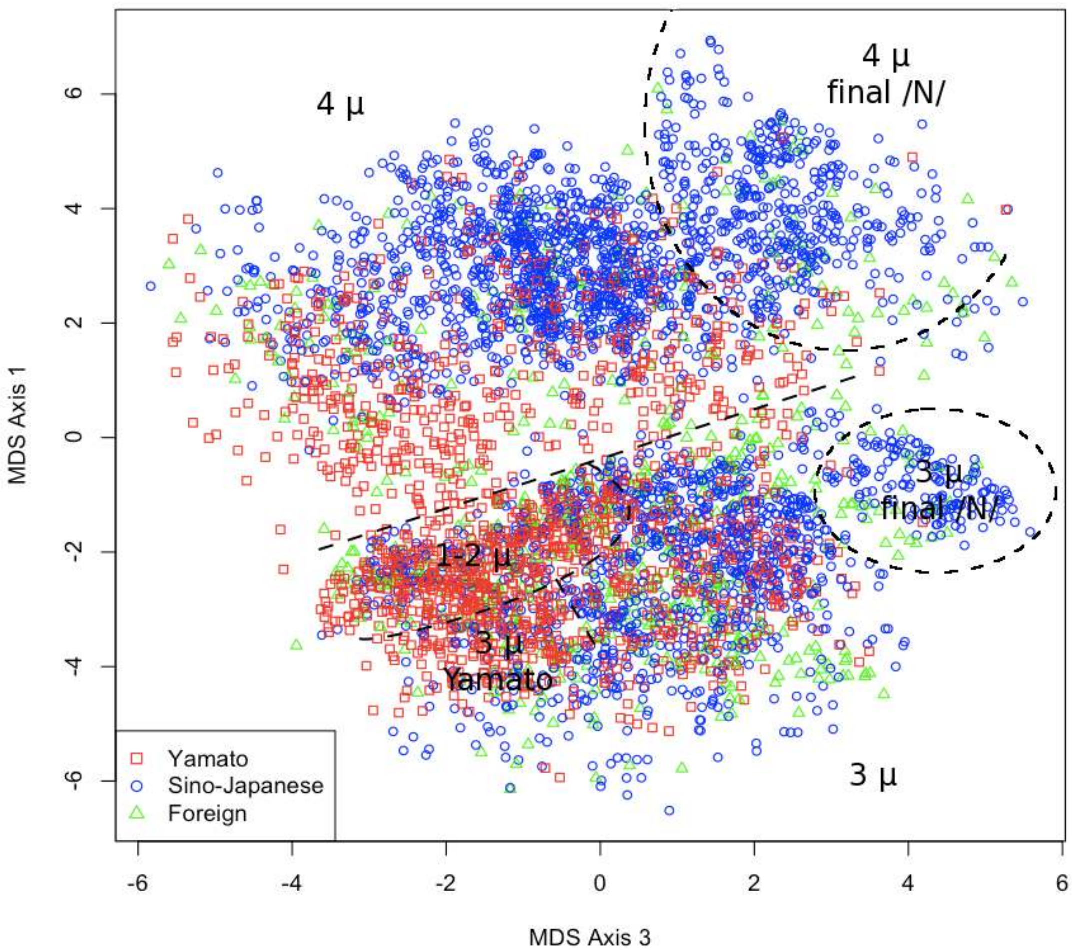


Figure 4.10: Scatterplot of MDS axes 1 and 3 from Figure 4.8, labeled with clusters corresponding to four-mora words, three-mora Yamato words, and three- and four-mora words with final /N/.

the middle two cells in the top row of Figure 4.8.) In Figure 4.9, the points fall into two distinct clusters, with the larger one across the top of the plot representing all words four moras in length,[19] and the smaller one on the bottom words that are less than four moras long.[20] Within the four-mora group, the Yamato and Sino-Japanese items are easily distinguishable, with little overlap between the two, while the Foreign items seem to be interspersed in smaller clusters throughout the larger four-mora group, but occurring mainly in the Yamato part. In the 1–2 μ and 3μ groups, on the other hand, it is difficult to distinguish Yamato and Sino-Japanese items according to their positions on MDS axes 1 and 2. It is somewhat easier to distinguish these items according to MDS axes 1 and 3, as shown in Figure 4.10, although even here there is still a significant amount of overlap. Again, Foreign items do not form a single large cluster on their own, but instead occur in smaller clusters interspersed throughout the MDS space. In Figure 4.10 it can also be seen that MDS axes 1 and 3 can be used to distinguish three- and four-mora words with final /N/ (most of which are members of the Sino-Japanese stratum) from the other items in the sample, since these words all occur in two clusters on the right side of the plot.

The results of the MDS analysis suggest that viewing lexical items as points within a similarity space can be effectively used by a language learner to classify lexical items by stratum, and thus to predict which phonological processes (such as *rendaku* or post-nasal voicing) a given lexical item will be governed by. While there is some overlap in the areas of similarity space which the Yamato, Sino-Japanese, and Foreign strata occupy, there is an overall global structure in the lexicon which the language learner can exploit in learning stratum classifications, since Yamato items tend to occur in certain

[19]This was determined by examining the individual lexical items associated with the points in the scatterplot.

[20]This particular clustering pattern, where words less than four moras in length are grouped closer together than four-mora words are, may be an artifact of the vector representation I use here, since all words that are less than four moras long would have a value of -1 for the first 26 elements of the vector, while four-mora words would instead have a value of 1 for at least some of these elements.

Table 4.10: Stratum classification training and test data

Stratum	Training set		Validation set		Test set		Totals	
Yamato	2,828	(33%)	2,731	(32%)	2,761	(33%)	8,320	(33%)
Sino-Japanese	4,721	(56%)	4,793	(57%)	4,795	(57%)	14,309	(56%)
Foreign	921	(11%)	946	(11%)	914	(11%)	2,781	(11%)
Totals	8,470		8,470		8,470		25,410	

regions of the similarity space, and Sino-Japanese items in other regions of the space. This suggests that the overall structure of the Japanese lexicon is more like Figure 4.7(b) than (a) or (c).

4.4.2 LDA classifier

It is possible to quantify the degree to which the clusters identified in the MDS analysis are distinguishable using another statistical method known as *linear discriminant analysis* (LDA; Venables & Ripley 2002). Given a set of observations with m features $x_0, x_1, \ldots x_m$, and a classification for each observation into one of n categories, LDA finds a set of $n-1$ linear combinations of the features (known as *linear discriminants*) which best divide the observations into the n categories. These linear discriminants can then be used to classify future observations. To construct an LDA-based classifier, all of the two, three, and four mora Yamato, Sino-Japanese, and Foreign words in `JMDICT` were randomly divided into three sets, a training set, a validation set, and a testing set, with an equal number of words in each set (Table 4.10). (The validation set was not used in the LDA analysis, but was used in the next section for training the connectionist classifier.) Linear discriminants for the three strata, Yamato, Sino-Japanese, and Foreign, for the items in the training set were found using the `lda` function in the `MASS` package for `R`. These linear discriminants were then used to predict the most likely classification for

Table 4.11: Confusion matrix for results from LDA classifier

	LDA prediction				
Stratum	Yamato	Sino-Japanese	Foreign	Accuracy	Completeness
Yamato	2,200	447	114	81%	80%
Sino-Japanese	253	4,413	129	85%	92%
Foreign	247	352	315	56%	34%

each of the items in the test set, and these predictions were compared with the original classifications given by the algorithm in section 4.3.2. Table 4.11 gives a confusion matrix for these results. In this table, the rows correspond to the orthographically-based stratum classifications, while the columns under "LDA prediction" give the predicted classifications using LDA on the feature representations of the words. For example, the Yamato row contains all of the 2761 Yamato items in the test set in Table 4.10; of these, 2200 were correctly classified by LDA as Yamato, 447 were incorrectly classified as Sino-Japanese, and 114 were incorrectly classified as Foreign. For each of the three strata, two measures of classification performance, *accuracy* and *completeness*, were computed as follows:

$$\text{accuracy} = \frac{\text{hits}}{\text{hits} + \text{false alarms}} \tag{4.8}$$

$$\text{completeness} = \frac{\text{hits}}{\text{hits} + \text{misses}}, \tag{4.9}$$

where *hits* are defined as the number of items in a given stratum classified correctly by LDA (for example, the 2200 Yamato items that were correctly classified); *misses* as the number of items in a stratum which were incorrectly classified (such as the 561 Yamato items which were classified as Sino-Japanese or Foreign); and *false alarms* the number of items with a given classification that are not actually members of that stratum (such as the 500 items classified as Yamato that are actually Sino-Japanese or Foreign). Accuracy measures how likely a given item classified as stratum S really is a member of S, while completeness measures how likely a member of S will be classified as S. It is evident from Table 4.11 that LDA performs best on classifying Sino-Japanese items (with 85% accuracy and 92% completeness), and worst on classifying Foreign items (with only 56% accuracy and 34% completeness). The low completeness score for Foreign items indicates that many Foreign items are difficult to distinguish phonologically from Yamato or Sino-Japanese items, which is what we would expect given that most Foreign items do not form distinct clusters in the MDS plots in Figure 4.8, but rather are distributed in smaller groups throughout the Yamato and Sino-Japanese clusters.

These results suggest that while the stratum classification task is difficult on the basis of phonological information only, it is not completely impossible, at least for Sino-Japanese and Yamato items. Ota's (2004) point still stands that lexical items can't be classified perfectly reliably on the basis of distributional constraints alone, such as whether a given item satisfies a constraint like *NT. However, if we combine multiple bits of distributional evidence, then the resulting classifier can be highly accurate, even if each cue on its own is only a weak predictor for a given stratum classification. Christiansen, Allen & Seidenberg (1998) call models like this a *multiple-cue integration* model, and note that this combining of multiple pieces of evidence to form a single decision is the type of task that CNs are quite good at (see also Christiansen, Dale & Reali In press). In this case, we might expect that a network trained to classify lexical items by stratum would potentially outperform an LDA-based classifier, since CNs are not limited to linear combinations of the phonological features in the lexical items, but can also consider nonlinear functions of these features, as well as the lexical frequencies of the items it is trained on, in making a classification.

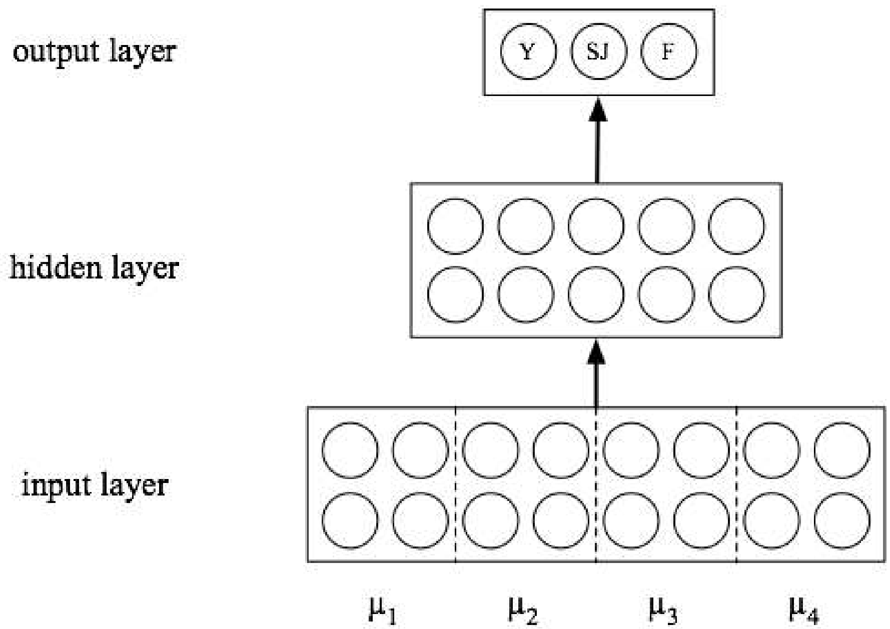


Figure 4.11: Network architecture for the stratum classification network

4.4.3 Feedforward network classifier

Ten feedforward networks were constructed with 104 input units (divided into four groups of 26 units, corresponding to the four moras of each input word), 50 tanh hidden units, and three softmax[21] output units indicating the predicted stratum of the input word: Yamato, Sino-Japanese, or Foreign (Figure 4.11). The networks were trained to associate the words in the training set from Table 4.10 with the appropriate stratum classification using the scaled momentum backpropagation algorithm[22] with cross-entropy error in the `Lens` network simulator (Rohde 2000), with a learning rate of 0.02 and momentum of 0.9. On-line training was used, with items probabilistically selected based

[21] Softmax units are logistic units which are normalized to sum to 1. These can be interpreted as posterior probabilities in a classification task (Sarle 2002).

[22] This algorithm "is similar to standard momentum descent with the exception that the pre-momentum weight step vector is bounded so that its length cannot exceed 1.0. After the momentum is added, the length of the resulting weight change vector can grow as high as 1 /(1 - momentum). This change allows stable behavior with much higher initial learning rates, resulting in less need to adjust the learning rate as training progresses." <http://tedlab.mit.edu/~dr/Lens/Commands/dougsMomentum.html>

Table 4.12: Confusion matrix for results from stratum classification network

	Network prediction				
Stratum	Yamato	Sino-Japanese	Foreign	Accuracy	Completeness
Yamato	2,344	327	90	87%	85%
Sino-Japanese	139	4,589	67	89%	96%
Foreign	223	259	432	73%	47%

on their log frequency in a database of text from Japanese novels (Kamermans 2008):

$$p = \frac{\log(\text{freq} + 1)}{\log \text{maxFreq}} \tag{4.10}$$

where maxFreq is the frequency of the highest frequency word in the sample. The networks' weights were initially set to random values in the range $[-0.1, 0.1]$, then the networks were trained for a total of 300,000 presentations of lexical items from the training set. Out of all of the networks that were trained, the network which performed the best on the validation set was chosen for analysis; this network's classification performance was assessed using the items in the test set. Neither the validation nor test set items were presented to any of the networks during training.

Table 4.12 presents a confusion matrix and accuracy and completeness measures for the stratum classification network. It is clear from this table that the network performs better overall than the LDA classifier (Table 4.11). While the accuracy and completeness measures for Yamato and Sino-Japanese items show some improvement when compared to LDA, the most noticeable improvement is with Foreign items, with 73% accuracy (as compared to 56% for the LDA classifier) and 47% completeness (as compared to 34% for LDA). This means that when the network classifies a given item as Foreign, we can be much more sure that this is the correct classification than we can with the LDA classifier, which almost half of the time misclassifies Yamato and Sino-Japanese items as Foreign. On the other hand, both classifiers misidentify over half of the Foreign items as either Yamato or Sino-Japanese, although the LDA classifier is much worse than the

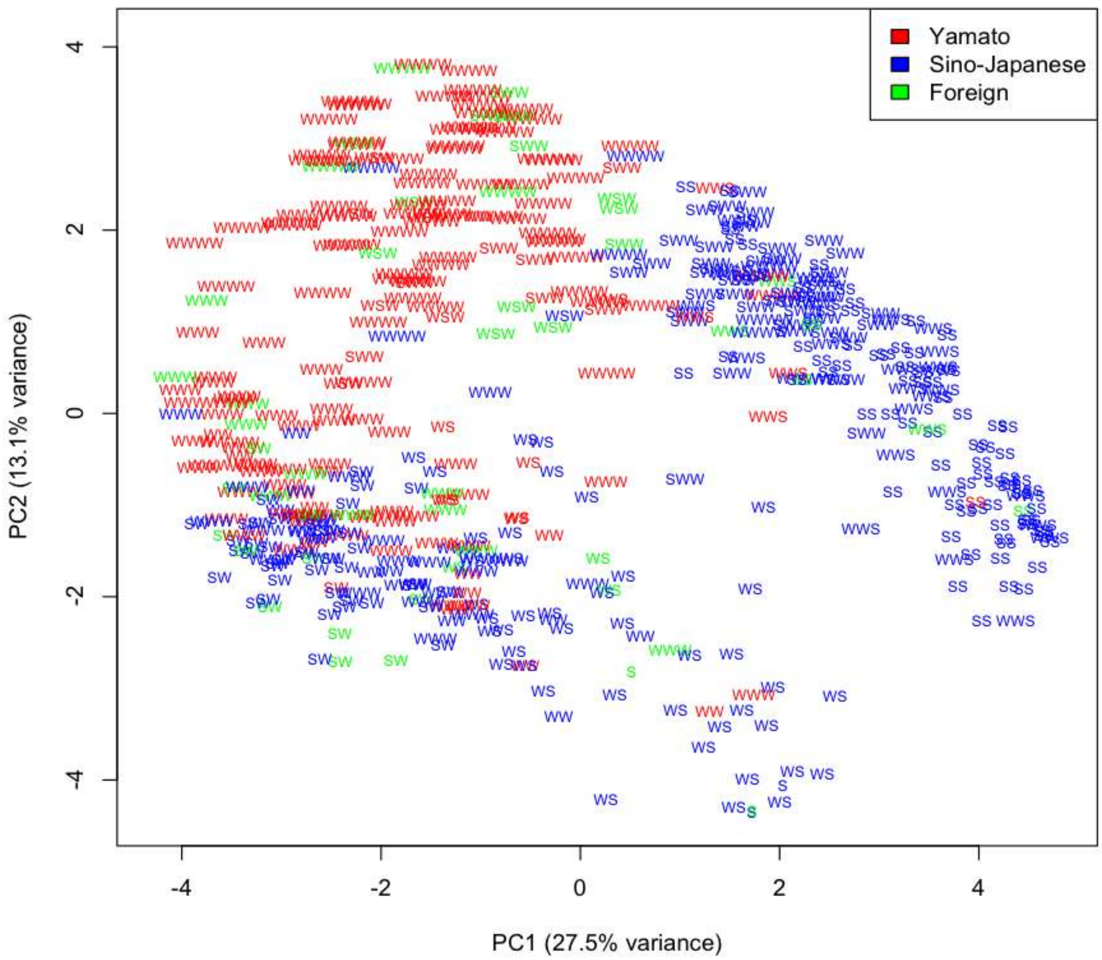


Figure 4.12: Principal component analysis of hidden unit activations in classification network. W=weak syllable; S=strong syllable.

network classifier in this respect.

The classifier network functions by mapping the feature representations on the input layer to internal representations on the hidden layer, and then identifying specific regions in the vector space defined by the hidden units with specific stratal classifications. This can be seen in Figure 4.12, which is a scatterplot showing the first two *principal components* (PCs) of the hidden unit activations of a random sample of the test items. Principal component analysis operates by identifying a small number of uncorrelated variables (the PCs) which account for the majority of the variation in a set of data, and

Table 4.13: Mean shortest distance from Foreign test items to Yamato, Sino-Japanese, and Foreign training items. Numbers in bold indicate the smallest distance in each row.

		Mean shortest distance		
Network prediction	# items	Yamato	Sino-Japanese	Foreign
Yamato	223	**2.36**	2.79	2.50
Sino-Japanese	259	2.65	**1.98**	2.30
Foreign	432	2.77	2.69	**2.23**

has been used in previous connectionist research, such as Elman (1991) and McClelland (1994), to examine the hidden unit representations that are generated in a trained CN. In Figure 4.12 I represent each test item as a sequence of weak (W) and strong (S) syllables, where W syllables have one mora, while S syllables have two. It is evident from the scatterplot that PC1, on the x-axis, encodes the weight of the final syllable in the word, with words ending in S syllables having higher values of PC1 than words ending in W syllables. Meanwhile, PC2, on the y-axis, encodes the length of the word in moras, with four-mora words having higher values of PC2 than three- and two-mora words. These two dimensions can be used to separate Yamato and Sino-Japanese items to a large extent. Yamato words are situated in a large cluster in the top left quadrant of the scatterplot (i.e. four-mora words ending in W syllables), while Sino-Japanese words occur in two smaller clusters in the bottom left and top right quadrants. However note that it is more difficult to distinguish Foreign items from the other two strata using the first two principal components, since Foreign items are scattered throughout the space defined by PC1 and PC2.

Of the Foreign items in the test set that are misclassified as Yamato or Sino-Japanese, it turns out that this misclassification is taking place because each of these items is relatively more similar to one of the Yamato or Sino-Japanese items in the training set than to any of the Foreign items in the training set. This is shown in Table 4.13, which compares the Euclidian distances from each of the Foreign test items to the items in the training set. For each Foreign test item, the distance to the closest Yamato, Sino-Japanese, and Foreign training item was recorded, and then the means of these distances were calculated, grouping the Foreign test items by the network's stratal classification for them. It can be seen from this table that all of the Foreign items in the test set classified as Yamato are, on average, closer to the Yamato items in the training set than to the Sino-Japanese or Foreign training items. Likewise, all of the Foreign test items classified as Sino-Japanese are closer to the Sino-Japanese training items than to the Yamato or Foreign training items. This shows that the classifier CN is performing the classification task on the basis of each test item's similarity to the lexical items the network was trained on (in addition to the classification in terms of syllable structure identified in the hidden unit PCA discussed above). Since Foreign items tend to be interspersed among the large-scale Yamato and Sino-Japanese clusters, it then becomes relatively more difficult to distinguish them from Yamato and Sino-Japanese on the basis of similarity to training items than it is to distinguish Yamato and Sino-Japanese words from each other. For a test word to be reliably identified as Foreign, it is necessary for there to be a Foreign item in the training set that is closer to it than any of the Sino-Japanese or Yamato items in the training set are.

4.4.4 Discussion

The performance of the LDA and network classifiers on the stratum classification task show that it is generally possible to distinguish members of the Yamato and Sino-Japanese strata on the basis of their phonological form. In order to do this, it is necessary to combine many different pieces of evidence about a word, such as the pattern of strong and weak syllables in the word, as well as the pattern of violation of the distributional constraints from Ito & Mester (1995), making the classification network an example of a multiple-cue integration model (Christiansen et al. 1998). The classification CN is not necessarily a psychologically realistic model of a Japanese speaker's lexical knowledge, since the task it was trained on (classifying lexical items by stratum) is not one that speakers need to perform in producing or understanding Japanese. Nevertheless, the classification CN shares an important property with TM networks, namely that it imposes a similarity metric on its input representations, and it exploits the similarity structure of the data it was trained on in order to perform the classification task. This means that we can think of lexical strata in connectionist models like TM as being emergent collections of smaller-grained phonological neighborhoods, just as in the stratum classification CN.

While Rice (1997) and Ota (2004) are correct in criticizing the Core-Periphery model for performing stratum classifications solely on the basis of phonological constraints such as *NT, it turns out that Ito & Mester's (1999) intuition regarding the Foreign stratum, namely that it is not a real stratum at all in the same sense as Yamato or Sino-Japanese, seems to correspond to the different kinds of distributions that the Yamato, Sino-Japanese, and Foreign items have in the similarity space of the Japanese lexicon. This is because, while the Yamato and Sino-Japanese strata form distinct, easily-identifiable clusters in this similarity space, Foreign items do not form a distinct cluster on their own, but are instead scattered in smaller groups throughout the Yamato and Sino-Japanese clusters. Likewise, Foreign items, unlike Sino-Japanese and Yamato items, do not form a coherent cluster in the hidden unit vector space of the stratum classification network. This means that the network found it difficult to map the feature values of the Foreign items to a well-defined region in the hidden unit space.

However, the representation of lexical strata used in Ito & Mester (1995) are quite different from the representations that develop in the trained stratum classification network. In Ito & Mester, the strata are the regions of overlapping domains of phonological constraints like *NT and *P, and there is no mechanism for representing the effects of phonological similarity on stratum classification. In the classification network, on the other hand, the network is not limited to the evidence from constraint satisfaction; it can also consider the similarity of test items to previously-learned lexical items in making a classification. In other words, using the terminology introduced in section 4.2, the stratum classification network imposes a metric topology on lexical items, which is a strictly finer topology than the partition topology of the Core-Periphery model. This allows the network to be able to learn that words like /koNbo/ or /teNba/, despite obeying all of the constraints governing the Yamato stratum, are actually members of different strata entirely.

4.5 Loanword adaptation network

In the stratum classification model that was just presented in the last section, we saw how the phonological representations learned by the network were organized on the basis of similarity. My claim is that lexical strata are emergent phenomena from this similarity structure that exists in the mental lexicon. In other words, strata are not explicitly specified characteristics of lexical entries (as they would be in theories where strata are represented using lexical features, like [+foreign]; Saciuk 1969) but instead are composed of smaller-scale phonological neighborhoods, where the entries within one of these neighborhoods are highly similar to each other phonologically, and where most of the members of a phonological neighborhood are in the same lexical stratum. In the previous section, we saw from the MDS plots of the training items that the members of the different strata were not distributed evenly throughout the space of possible phonological forms, but in fact were clustered in several subareas of this space. The network learned the stratum classification task by mapping these smaller neighborhoods to specific regions in the hidden unit space which corresponded to the classification categories (Yamato, Sino-Japanese, or Foreign).

Since the effects of phonological neighborhoods are also implicated in the changes in loanword adaptation patterns over time, as discussed in Chapter 2, it is natural then to consider a connectionist model of loanword adaptation in which these similarity effects arise as a result of the similarity structure of the lexical representations used by the network. Specifically, what I will develop in this section is a set of CNs which are first trained on various subsets of Japanese lexical items in order to learn Japanese phonological patterns and constraints. These networks are then presented several loanwords to see how they behave with respect to non-native phonotactic patterns. The prediction is that the adaptations produced by the networks will be effected by the phonological similarity of the loanwords to the various phonological neighborhoods already represented in the networks. In this experiment, I will be focusing specifically on loanwords with voiced geminates, since there are a relatively large number of them in the data I collected from Arakawa (1977), and because they show phonological neighborhood effects based on the place of articulation of the final consonant.[23]

4.5.1 Methods

In order to look at what kinds of adaptations a network will produce when performing an adaptation of a new lexical item, it is necessary to first train it on a subset of Japanese words so that it can learn phonological processes such as nasal place assimilation after

[23]Unfortunately, while the set of loans with source coronals before /i/ would be an even better demonstration of the effects of phonological neighborhoods in a CN, since there are many finely-grained neighborhoods through which the change from the TI→ČI to TI→TI adaptation strategy swept through, there are not enough members of many of these neighborhoods in the data I have collected to provide enough examples to train a network with.

Table 4.14: Mora alignment on input layer of adaptation network

Two-mora words:	μ_1	Ø	μ_2
Three-mora words:	μ_1	μ_2	μ_3

/N/. This was done as follows. All of the two- and three-mora Yamato, Sino-Japanese, and Foreign words from JMDICT were randomly divided into a training and test set of equal size. This particular subset of the Japanese vocabulary was chosen for training so that I could examine how the network will behave with loanwords with voiced geminates, which require a minimum of two syllables and three moras to be represented in Japanese (for example, ベッド /beQ.do/ 'bed', made up of the three moras /be/, /Q/, and /do/). When the training words are presented to the network, the moras are aligned so that the the word-final mora is always presented in the μ_3 slot, while the word-initial mora is presented in the μ_1 slot. Meanwhile, μ_2 either contains the word-medial mora (if the word is three moras in length) or is empty, with all place and manner features set to -1, as shown in Table 4.14.

The network architecture is similar to the attractor network used in the first simulation in Harm & Seidenberg (1999), as shown in Figure 4.13. There are two layers of units: 84 feature units, divided into three groups of 28 units (26 feature units and two prosody units), representing the moras in the word, and using the feature representation from section 4.3.1; and 50 cleanup units, with bidirectional connections between the two sets of units. In addition there are recurrent connections within the feature layer: each feature unit is connected to all other feature units, including itself. Each feature unit's self-connection has a fixed weight of 0.75, while the weights to other feature units were trainable.

The goal of the network during training is to learn how to reproduce the input words presented using the feature units. The presentation of each training item took place over

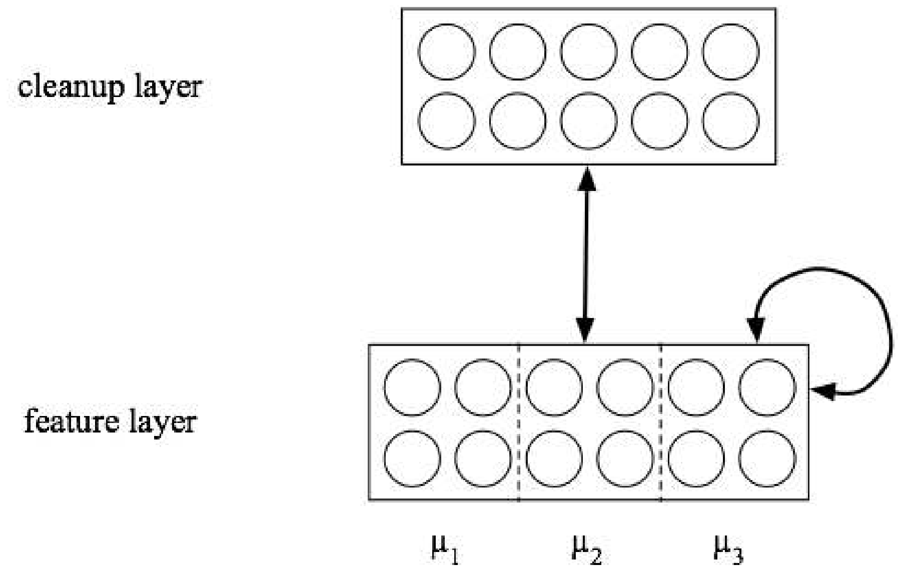


Figure 4.13: Architecture of the adaptation network

six "ticks" (discrete time units) on the network. On the first tick all of the feature and cleanup units in the network are set to the value 0. Then on the second tick, a word was chosen from the training sample with a probability p proportional to its log frequency in Kamermans (2008):

$$p = \frac{\log(freq+1)}{\log maxFreq} \tag{4.11}$$

where *maxFreq* is the frequency of the highest frequency word in the sample. The word was converted into the feature representation from section 4.3.1 and presented on the feature units of the network. The network was then run for four more ticks, with each unit u_i computing its activation a_i based on the weighted activations $w_{ij}a_j$ on the previous tick of all of the units u_j connected to it:

$$a_i = \tanh(\sum_j w_{ij}a_j + \theta_i). \tag{4.12}$$

Then on ticks four, five, and six, the activations on the phoneme units were compared to the original input word, and the sum-squared error on these two ticks were used to train the weights in the network using the *backpropagation through time* method (Williams & Peng 1990). Since each feature unit has a connection to itself with a fixed weight of

0.75, if the other weights in the network are close to 0 (as they are in the initial state before training), then what would happen on each following tick is that the activation on the feature units would gradually decrease. Thus the network must train the other recurrent weights on the feature layer and the bidirectional weights between the feature and hidden layers in order to successfully reproduce and maintain the input pattern on the last three ticks.

There were four different networks with this architecture that were trained. The training and test data for Network 1 was generated by identifying all of the two- and three-mora words in `JMDICT` in the Yamato, Sino-Japanese, and Foreign strata (excluding a set of 15 loanwords with voiced geminates, listed in Table 4.17 in the Results section below). These words were randomly divided into two sets of equal size, a training set and a test set. This training condition is meant to represent the lexical knowledge that a modern-day Japanese speaker would have. Network 2 was trained and tested using Yamato and Sino-Japanese items only, with no Foreign items present. This is meant to represent the lexical knowledge that an early 19th-century Japanese speaker would have, when there were few loanwords in common use.[24]

The training data for Networks 3 and 4 were constructed to test a couple of possible factors explaining why loanwords with final /d/ were the first to be attested as voiced geminates in Japanese, as discussed in section 2.3.4 of Chapter 2. Network 3 was trained on a random sample similar to Network 2 consisting of Yamato and Sino-Japanese items

[24] I made two simplifying assumptions in the construction of the training and test data for Network 2: one, that lexical frequencies for most Yamato and Sino-Japanese items have not changed very much from the early 19th century until the present day, when the frequency data in Kamermans (2008) was compiled; and two, that the relative proportions of Yamato and Sino-Japanese items in the Japanese lexicon at that time were comparable to those of today. Both assumptions are rather questionable, since many Sino-Japanese items were coined beginning in the Meiji period, in the late 19th century (Shibatani 1999), and so the frequencies for many of the Sino-Japanese items in the training sample would have been much lower in the early 19th century than they are in the present day, if they were even attested at that time. A better (but considerably more difficult) way of constructing the training set for Network 2 would be to take a sample of lexical items from early 19th-century texts, and estimate their frequencies from these texts. I leave this for future research.

only, but with the addition of the single loanword ベッド /beddo/ 'bed'. This was done to see what the effect would be on voiced geminate adaptation patterns on a network which had only been exposed to a singe example of a word with a voiced geminate. The word /beddo/ was chosen in particular because it was one of the few loanwords to be attested with a voiced geminate before 1870 (Arakawa 1977). The devoiced form /betto/ was also attested at that time, so in order to represent this variation, both forms were added to the training set for Network 3.[25] The training set for Network 4, on the other hand, included no loanwords, just like Network 2; however, for all of the verbs selected in the training sample, the network was trained on both the infinitive and past tense forms of the verb (if the past tense form was two or three moras in length). There were 133 such verbs in the training set. For example, the verb 待つ /matɯ/ 'to wait' was presented to the network half of the time in its infinitive form [matsɯ] and the other half of the time in its past tense form [matta]. This results in a larger number of training items with geminate /tt/ as compared to the training sets for the other three networks, which used the infinitive forms only of any verbs that were present (Figures 4.14 and 4.15).

Table 4.15 summarizes the relative proportions of Yamato, Sino-Japanese, and Foreign items in the training and testing samples for all of the networks that were trained.

During training, each of the input feature values on tick 1 had a small amount of Gaussian noise (with a mean of 0 and a standard deviation of 0.05) added. This results in the networks learning a smoother decision function than they would with non-noisy inputs, which helps reduce the likelihood of overfitting and improves the generalization

[25]The training data for Network 3 has the same issues with regards to frequency estimates as does the training data for Network 2, as discussed in footnote 24. The raw frequency for /beddo/ 'bed' was taken from the frequency counts in Kamermans (2008), which almost certainly overestimates the actual frequency that the loanword had when it was first borrowed. There were three attestations of *bed* dating from before 1870 in Arakawa (1977); of these, one had a voiced geminate (/beddo/) and the other two were devoiced (/betto/). To represent this variation, both forms were included in the training set for Network 3, with the /betto/ form having a raw frequency of $5476\frac{2}{3}$, twice that of /beddo/, which was $2738\frac{1}{3}$. The frequencies of both forms sum to 8215, which is the frequency of modern-day /beddo/ given in Kamermans (2008).

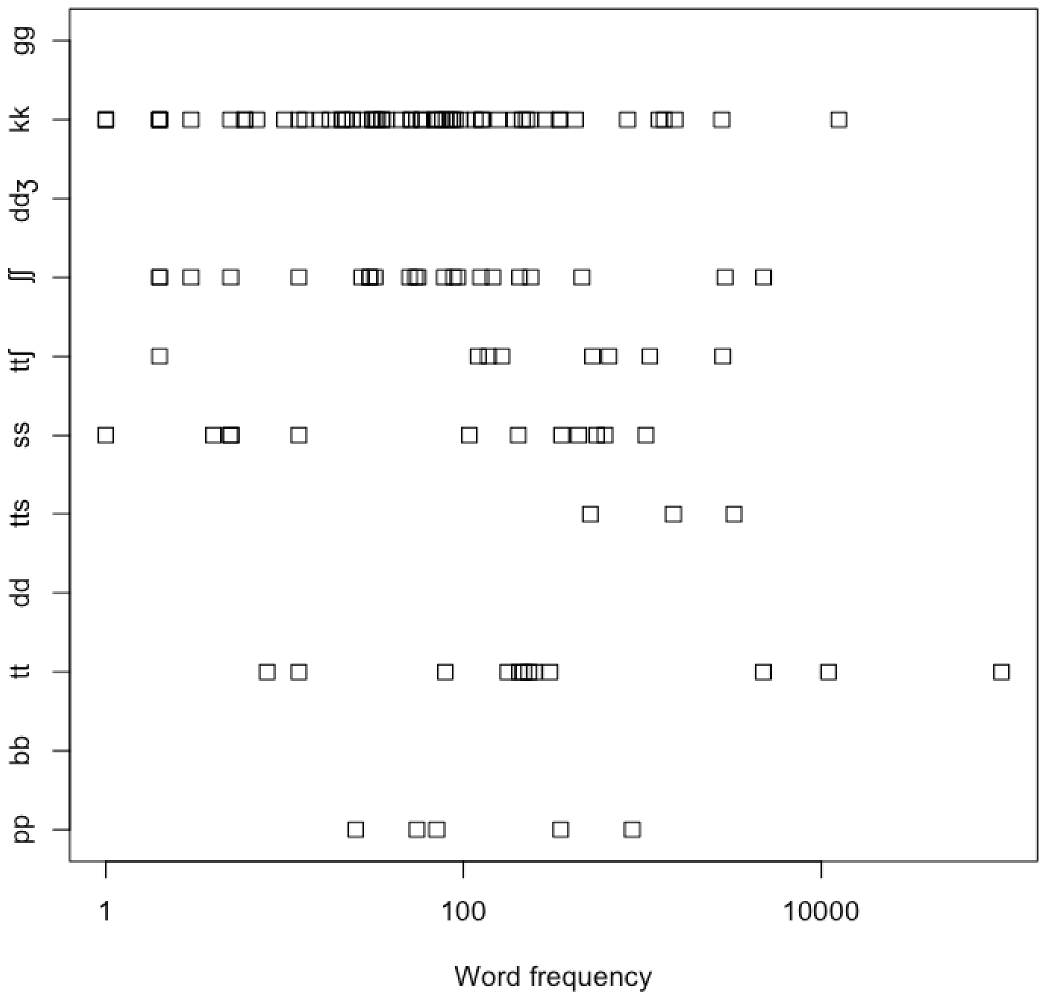


Figure 4.14: Frequencies of words containing geminate obstruents in the training set for Network 2. Each box corresponds to a single lexical item. Only the infinitive forms of verbs are included in this training set.

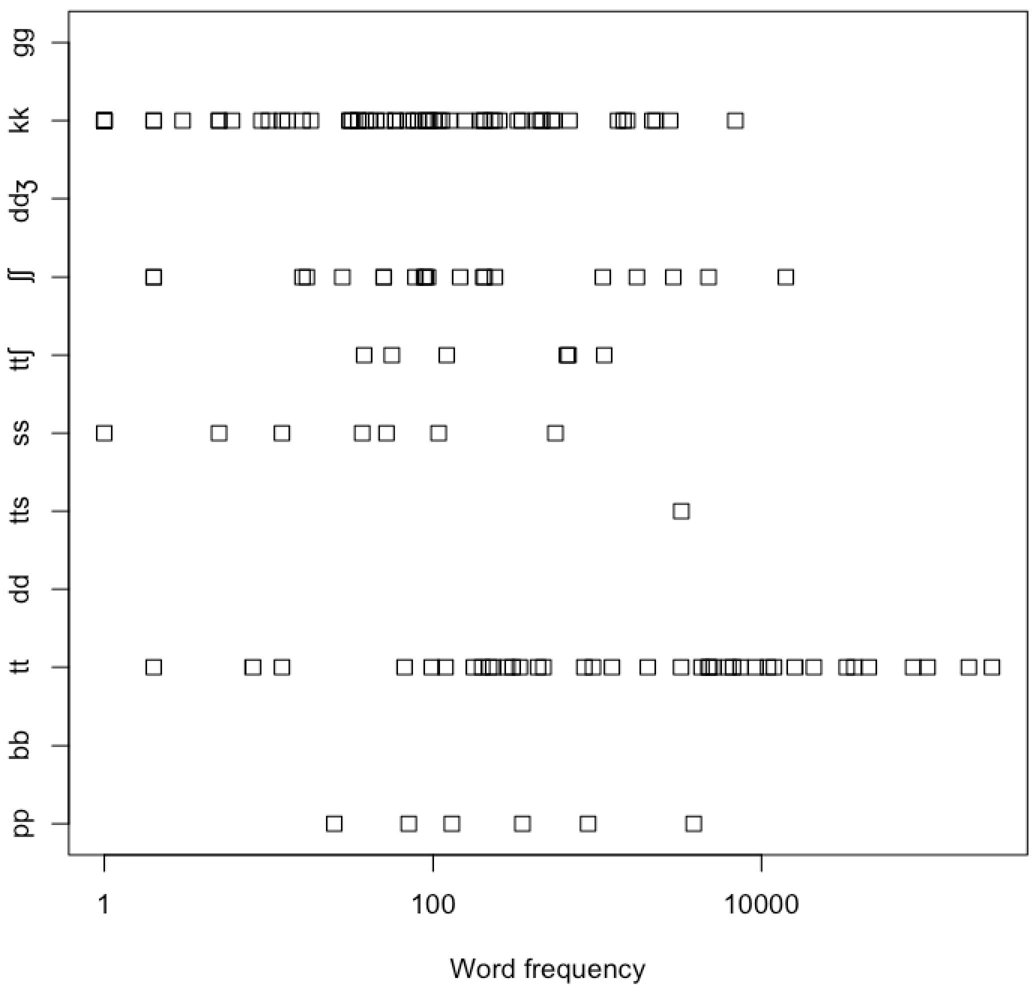


Figure 4.15: Frequencies of words containing geminate obstruents in the training set for Network 4. Past tense forms of verbs are included in this training set, resulting in many more examples of geminate /tt/ than in Figure 4.14.

Table 4.15: Training and test data for adaptation networks

	Stratum	Training set		Testing set	
Network 1	Yamato	1,838	(34%)	1,837	(34%)
	Sino-Japanese	2,862	(54%)	2,831	(53%)
	Foreign	646	(12%)	678	(13%)
	Totals	5,346		5,346	
Network 2	Yamato	1,843	(39%)	1,853	(39%)
	Sino-Japanese	2,876	(61%)	2,866	(61%)
	Foreign	0		0	
	Totals	4,719		4,719	
Network 3	Yamato	1,828	(39%)	1,868	(40%)
	Sino-Japanese	2,891	(61%)	2,851	(60%)
	Foreign	1		0	
	Totals	4,720		4,719	
Network 4	Yamato	1,960	(40%)	1,869	(40%)
	Sino-Japanese	2,892	(60%)	2,850	(60%)
	Foreign	0		0	
	Totals	4,852		4,719	

Table 4.16: Adaptation network performance on training and testing items

Network	Training set	Test set
1	95%	92%
2	97%	94%
3	95%	92%
4	96%	92%

capabilities of the networks (Reed & Marks 1999). To further improve generalization, during training the input on the first tick for each of the μ_2 nucleus feature units had a small probability ($p = 0.01$) of being changed to a random value in the range $[-1,1]$ (the target values for these units remained the same, however). All of the networks were trained using on-line training, meaning that the networks' weights were updated immediately after the presentation of each word selected from the training sample. A learning rate of 0.005 was used, and training was stopped after $300,000$ trials, at which point the networks' performance was assessed.

4.5.2 Results

The accuracy of the networks at the repetition task was measured by presenting each of the items in their respective training and testing sets one at a time, and then comparing the activation on each feature unit on ticks 3, 4, and 5 with the input activation that was presented on tick 1. If, for all feature units, the activation of each feature unit was within 0.25 of the target activation, then the network was considered to have successfully reproduced the input pattern. Table 4.16 shows the results of assessing each of the networks using this metric. The high accuracy on both the training and testing samples shows that all of the networks successfully learned the repetition task, since they were able to generalize the task to the words from the testing sample, which had not been exposed to the networks at all during training.

Table 4.17: Test words for adaptation networks

	/bb/			/dd/		
/a_/	カッブ	kabbɯ	‘cub’			
/i_/	ジッブ	ʤibbɯ	‘jib’	キッド	kiddo	‘kid’
/ɯ_/				ウッド	uddo	‘wood’
/e_/	ウェッブ	webbɯ	‘web’	デッド	deddo	‘dead’
/o_/	ボッブ	bobbɯ	‘bob’	ゴッド	goddo	‘god’
	/dʤ/			**/gg/**		
/a_/	バッジ	badʤi	‘badge’	マッグ	maggɯ	‘mug’
/i_/	リッジ	ridʤi	‘ridge’	ジッグ	ʤiggɯ	‘jig’
/ɯ_/						
/e_/	エッジ	edʤi	‘edge’			
/o_/	ロッジ	rodʤi	‘lodge’	ホッグ	hoggɯ	‘hog’

To look at how each network adapts loanwords with voiced geminates, I then presented an additional set of 15 loanwords from JMDICT each containing a voiced geminate /b/, /d/, /ʤ/, or /g/ (Table 4.17). None of these words were present in the training or testing sets of any of the networks. I tried as best as possible to find examples of voiced geminates occurring after each of the five vowels of Japanese; however there are some combinations, such as /ɯbb/, which are unattested in JMDICT.[26] All of the words in Table 4.17 are attested with geminates in JMDICT, except for the four words in the /b/ column. Of these, カブ /kabɯ/ ‘cub’ and ジブ /ʤibɯ/ ‘jib’ are attested only with singleton /b/ in JMDICT, while ウェ(ッ)ブ /we(b)bɯ/ ‘web’ and ボ(ッ)ブ /bo(b)bɯ/ ‘bob’ are attested with both singleton and geminate /b/.

All of the test words in Table 4.17 were presented to the four networks ten times each, then the average activations of the feature units on the last tick of each presentation were recorded and compared to three possible outcomes: voiced geminate (e.g. /kiddo/ for *kid*); devoiced geminate (/kitto/); and nasal-voiced stop cluster (/kiɴdo/). Of these

[26]The vowel following the geminate in each loanword is epenthetic, and is determined by the place of articulation of the geminate: /ɯ/ for /bb/ and /gg/, /o/ for /dd/, and /i/ for /dʤ/.

outcomes, the first two are attested historically as possible adaptations of voiced geminates, as discussed in section 2.3.4 of Chapter 2. The third outcome, the nasal-voiced stop cluster, is never attested as a possible adaptation, though Ichikawa (1930: 182; quoted in Lovins 1975) has an interesting anecdote concerning this possibility:

> The occasional use of voiceless for voiced plosives is to be noted, especially after short vowels. Exx. [kitto] ('kid'), [betto] ('bed'), [hettoraito] ('headlight'), [opera-pakku] ('opera-bag'). In these cases it is the difficulty of pronouncing short vowels before long voiced plosives which makes us prefer voiceless sounds. A maid, when asked to pronounce [beddo], said [betto] and, on being told to keep the *d-sound*, only succeeded in doing so by pronouncing it [bendo].

As Lovins (1975) points out, a nasal-voiced stop cluster adaptation would preserve both the voicing of the source word stop and the closed nature of the final syllable of the source word. However, it has never been used as a possible adaptation for voiced geminates, presumably because borrowers are aware that there is no nasal present in the source word, and creating a nasal-stop sequence is perceived as being less faithful to the source word than the other possible adaptations.

Table 4.18 shows the results of presenting the words in Table 4.17 to Network 1. Here the closest adaptation is determined by interpreting the activation on the feature units as an 84-element vector, and computing the distance from that vector to each of the expected feature values for each of the three possible adaptations enumerated above:

$$\text{dist} = \sqrt{\sum_i (a_i - t_i)^2}, \tag{4.13}$$

where a_i are the values of each of the feature units, and t_i the target feature values for the adaptation that the current activation is being compared with. The adaptation with the smallest distance from the activation for each word was considered the closest adaptation, while the adaptation with the next smallest distance was considered the next closest. In one sense, Network 1 has learned that voiced geminates are phonologically valid sequences in modern Japanese, since the closest adaptation pattern for every word in Table 4.18 is VCC. At the same time, though, the acceptability of the VCC adaptation pattern varies by place of articulation, as shown by the varying adaptation distances and average activations of the [voi] feature of μ_2. Words with geminate /d/ and /g/ have lower distances from VCC, and the higher mean activations for [voi], than words with geminate /b/. This shows that Network 1 has also learned that /bb/ is less acceptable than /dd/ or /gg/ in Japanese.[27]

Network 2, which was trained on Yamato and Sino-Japanese items only, shows an overall reduction in the acceptability of voiced geminates, as shown in Table 4.19. Unlike with Network 1, there are no test items for which Network 2 unambiguously produces a VCC output. Instead, the network produces VC̥C̥ adaptations for words with /bb/, /dʤ/, or /gg/, as well as two of the words with /dd/ (/kiddo/ 'kid' and /ɯddo/ 'wood'), and VNC adaptations for the other two words with /dd/ (/deddo/ 'dead' and /goddo/ 'god'). These two adaptations are a result of the network trying to satisfy the *DD constraint (avoid producing a voiced geminate) while also preserving μ_2 as a consonant (ruling out a repair via deletion of μ_2, or changing μ_2 into a vowel). This can be done by either devoicing the input segments, creating a voiceless geminate, or by changing μ_2 into a moraic nasal, creating a nasal-voiced stop sequence. This latter option also preserves the voiced quality of the input segments, which Network 2 seems to value more than preserving the value of the [nas] and [son] features. This behavior is

[27] Note that in the results for all of the networks, the test words with geminate /dʤ/ tend to pattern with the words with /bb/, instead of with /dd/ and /gg/, as we would expect based on the attested adaptation data in Japanese. This seems to be because of the relative rarity of loanwords with /ʤ/, and because the feature representation does not allow the networks to discover that /ʤ/ should pattern with /d/ in some contexts (for example, underlying /di/ → [ʤi]), since /d/ and /ʤ/ differ in the values for the [ant], [cont_2], and [strid] features.

Table 4.18: Adaptation patterns of voiced geminates produced by Network 1. VCC=voiced geminate; VC̥C̥=devoiced geminate; VNC=nasal-voiced stop cluster. Numbers in parentheses indicate the Euclidean vector distance from the adaptation pattern to the actual feature unit activation in Network 1.

		μ_2 mean activation							
	Word	[son]	[cons]	[voi]	[nas]	Closest adaptation		Next closest adaptation	
/b/	/kabbɯ/ ‘cub’	−0.702	0.950	0.538	−0.749	VCC	(0.646)	VC̥C̥	(1.599)
	/ʤibbɯ/ ‘jib’	−0.552	0.943	0.775	−0.472	VCC	(0.765)	VC̥C̥	(1.924)
	/webbɯ/ ‘web’	−0.769	0.924	0.630	−0.834	VCC	(0.521)	VC̥C̥	(1.670)
	/bobbɯ/ ‘bob’	−0.662	0.956	0.790	−0.586	VCC	(0.636)	VC̥C̥	(1.889)
/d/	/kiddo/ ‘kid’	−0.867	0.862	0.741	−0.992	VCC	(0.387)	VC̥C̥	(1.764)
	/ɯddo/ ‘wood’	−0.922	0.860	0.728	−0.989	VCC	(0.457)	VC̥C̥	(1.767)
	/deddo/ ‘dead’	−0.829	0.890	0.839	−0.971	VCC	(0.368)	VC̥C̥	(1.869)
	/goddo/ ‘god’	−0.842	0.894	0.834	−0.966	VCC	(0.340)	VC̥C̥	(1.858)
/ʤ/	/badʤi/ ‘badge’	−0.916	0.984	0.540	−0.990	VCC	(0.574)	VC̥C̥	(1.578)
	/ɾidʤi/ ‘ridge’	−0.888	0.974	0.703	−0.989	VCC	(0.534)	VC̥C̥	(1.759)
	/edʤi/ ‘edge’	−0.884	0.985	0.683	−0.985	VCC	(0.493)	VC̥C̥	(1.724)
	/ɾodʤi/ ‘lodge’	−0.916	0.980	0.615	−0.990	VCC	(0.554)	VC̥C̥	(1.663)
/g/	/maggɯ/ ‘mug’	−0.876	0.870	0.619	−0.991	VCC	(0.470)	VC̥C̥	(1.642)
	/ʤiggɯ/ ‘jig’	−0.860	0.816	0.745	−0.987	VCC	(0.416)	VC̥C̥	(1.776)
	/hoggɯ/ ‘hog’	−0.880	0.868	0.629	−0.982	VCC	(0.461)	VC̥C̥	(1.651)

quite different from that of historical Japanese speakers, since VNC is never attested as a possible adaptation for a voiced geminate.

The training data for Networks 3 and 4 were constructed to test the effects of exposure to a single loanword, /beddo/ 'bed', and training on both the infinitive and past tense forms of verbs, respectively, on the likelihood of a geminate /d/ being produced by the network. Table 4.20 shows the adaptations produced by Network 3 for each of the test words. Here we can see that the presence of /beddo/ in the training set had a significant effect on this network's adaptation patterns. All of the words with final /d/, as well as one of the /g/ words, /ʤiggɯ/ 'jig', are produced with the VCC adaptation, while words with final /ʤ/ and /b/ are instead produced with the VC̥C̥ adaptation. None of the test items are produced with the historically unattested VNC adaptation. Among the words with /d/, the word /deddo/ 'dead', which is the most phonologically similar to the training item /beddo/, shows the largest mean activation of the [voi] feature.

Network 4 behaved somewhat similarly to Network 2, in that most of the test items are adapted as VC̥C̥. The only exceptions are three VNC adaptations for /deddo/ 'dead', /goddo/ 'god', and /edʤi/ 'edge', and one VCC adaptation for /kiddo/ 'kid'.

Table 4.22 summarizes the most common adaptation pattern for the four networks for each of the test loanwords.

4.5.3 Discussion

The results of the four networks in producing loanword adaptations show that at least some of the frequency and phonological neighborhood effects seen in historical adaptation patterns arise naturally as a result of the structure and behavior of the TM model

Table 4.19: Adaptation patterns of voiced geminates produced by Network 2

		μ_2 mean activation							
	Word	[son]	[cons]	[voi]	[nas]	Closest adaptation		Next closest adaptation	
/b/	/kabbɯ/ ‘cub’	−0.503	0.879	−0.408	−0.925	VC̥C̥	(0.814)	VCC	(1.525)
	/ʤibbɯ/ ‘jib’	−0.834	0.906	−0.801	−0.956	VC̥C̥	(0.322)	VCC	(1.820)
	/webbɯ/ ‘web’	−0.673	0.891	−0.604	−0.937	VC̥C̥	(0.554)	VCC	(1.652)
	/bobbɯ/ ‘bob’	−0.866	0.904	−0.822	−0.985	VC̥C̥	(0.320)	VCC	(1.841)
/d/	/kiddo/ ‘kid’	−0.139	0.923	−0.110	−0.895	VC̥C̥	(1.251)	VCC	(1.434)
	/ɯddo/ ‘wood’	−0.226	0.887	−0.156	−0.901	VC̥C̥	(1.161)	VCC	(1.417)
	/deddo/ ‘dead’	0.667	0.954	0.744	0.052	VNC	(1.043)	VCC	(1.997)
	/goddo/ ‘god’	0.472	0.926	0.589	−0.450	VNC	(1.610)	VCC	(1.658)
/ʤ/	/badʤi/ ‘badge’	−0.965	0.991	−0.953	−0.996	VC̥C̥	(0.204)	VCC	(1.963)
	/ridʤi/ ‘ridge’	−0.935	0.990	−0.915	−0.986	VC̥C̥	(0.242)	VCC	(1.928)
	/edʤi/ ‘edge’	−0.929	0.992	−0.913	−0.986	VC̥C̥	(0.272)	VCC	(1.930)
	/rodʤi/ ‘lodge’	−0.910	0.990	−0.895	−0.990	VC̥C̥	(0.295)	VCC	(1.915)
/g/	/maggɯ/ ‘mug’	−0.755	0.875	−0.724	−0.993	VC̥C̥	(0.403)	VCC	(1.749)
	/ʤiggɯ/ ‘jig’	−0.372	0.890	−0.286	−0.697	VC̥C̥	(1.013)	VCC	(1.510)
	/hoggɯ/ ‘hog’	−0.331	0.853	−0.273	−0.939	VC̥C̥	(1.008)	VCC	(1.458)

Table 4.20: Adaptation patterns of voiced geminates produced by Network 3

	Word	μ_2 mean activation				Closest adaptation		Next closest adaptation	
		[son]	[cons]	[voi]	[nas]				
/b/	/kabbɯ/ ‘cub’	−0.886	0.942	−0.727	−0.955	VC̥C̥	(0.384)	VCC	(1.747)
	/ʤibbɯ/ ‘jib’	−0.540	0.959	−0.172	−0.350	VC̥C̥	(1.164)	VCC	(1.465)
	/webbɯ/ ‘web’	−0.858	0.932	−0.679	−0.963	VC̥C̥	(0.413)	VCC	(1.699)
	/bobbɯ/ ‘bob’	−0.843	0.930	−0.416	−0.904	VC̥C̥	(0.667)	VCC	(1.453)
/d/	/kiddo/ ‘kid’	−0.546	0.950	0.236	−0.821	VCC	(0.937)	VC̥C̥	(1.337)
	/ɯddo/ ‘wood’	−0.530	0.953	0.192	−0.768	VCC	(0.983)	VC̥C̥	(1.309)
	/deddo/ ‘dead’	−0.275	0.952	0.692	−0.646	VCC	(0.894)	VC̥C̥	(1.889)
	/goddo/ ‘god’	−0.628	0.960	0.371	−0.780	VCC	(0.795)	VC̥C̥	(1.445)
/ʤ/	/badʤi/ ‘badge’	−0.963	0.992	−0.843	−0.986	VC̥C̥	(0.405)	VCC	(1.880)
	/ɾidʤi/ ‘ridge’	−0.921	0.990	−0.891	−0.981	VC̥C̥	(0.324)	VCC	(1.915)
	/edʤi/ ‘edge’	−0.943	0.992	−0.836	−0.989	VC̥C̥	(0.390)	VCC	(1.870)
	/ɾodʤi/ ‘lodge’	−0.909	0.988	−0.816	−0.976	VC̥C̥	(0.478)	VCC	(1.870)
/g/	/maggɯ/ ‘mug’	−0.796	0.882	−0.524	−0.964	VC̥C̥	(0.553)	VCC	(1.551)
	/ʤiggɯ/ ‘jig’	−0.379	0.889	0.036	−0.693	VCC	(1.228)	VC̥C̥	(1.259)
	/hoggɯ/ ‘hog’	−0.397	0.860	−0.075	−0.898	VC̥C̥	(1.125)	VCC	(1.296)

Table 4.21: Adaptation patterns of voiced geminates produced by Network 4

	Word	μ_2 mean activation				Closest adaptation		Next closest adaptation	
		[son]	[cons]	[voi]	[nas]				
/b/	/kabbɯ/ ‘cub’	−0.766	0.885	−0.752	−0.972	VC̥C̥	(0.429)	VCC	(1.786)
	/ʤibbɯ/ ‘jib’	−0.749	0.907	−0.727	−0.933	VC̥C̥	(0.442)	VCC	(1.763)
	/webbɯ/ ‘web’	−0.711	0.914	−0.737	−0.876	VC̥C̥	(0.477)	VCC	(1.783)
	/bobbɯ/ ‘bob’	−0.411	0.870	−0.378	−0.834	VC̥C̥	(0.921)	VCC	(1.564)
/d/	/kiddo/ ‘kid’	0.101	0.970	0.082	−0.130	VCC	(1.699)	VNC	(1.714)
	/ɯddo/ ‘wood’	−0.148	0.956	−0.152	−0.275	VC̥C̥	(1.411)	VCC	(1.669)
	/deddo/ ‘dead’	0.945	0.955	0.947	0.996	VNC	(0.125)	VCC	(2.789)
	/goddo/ ‘god’	0.840	0.950	0.844	0.911	VNC	(0.265)	VCC	(2.661)
/ʤ/	/badʤi/ ‘badge’	−0.535	0.988	−0.504	−0.565	VC̥C̥	(1.071)	VCC	(1.798)
	/ɾidʤi/ ‘ridge’	−0.910	0.993	−0.902	−0.964	VC̥C̥	(0.241)	VCC	(1.915)
	/edʤi/ ‘edge’	0.090	0.990	0.052	0.190	VNC	(1.791)	VCC	(2.220)
	/ɾodʤi/ ‘lodge’	−0.830	0.984	−0.792	−0.965	VC̥C̥	(0.585)	VCC	(1.875)
/g/	/maggɯ/ ‘mug’	−0.524	0.815	−0.518	−0.942	VC̥C̥	(0.722)	VCC	(1.617)
	/ʤiggɯ/ ‘jig’	−0.273	0.830	−0.274	−0.870	VC̥C̥	(1.067)	VCC	(1.505)
	/hoggɯ/ ‘hog’	−0.002	0.575	−0.021	−0.928	VC̥C̥	(1.476)	VCC	(1.525)

Table 4.22: Most common adaptation patterns produced by the four networks for each test loanword. Numbers in parentheses indicate the mean Euclidean distance from the adaptation pattern to the output unit activation produced by each network.

	Word	Network 1		Network 2		Network 3		Network 4	
/b/	/kabbɯ/ ‘cub’	VCC	(0.646)	VC̥C̥	(0.814)	VC̥C̥	(0.384)	VC̥C̥	(0.429)
	/ʤibbɯ/ ‘jib’	VCC	(0.765)	VC̥C̥	(0.322)	VC̥C̥	(1.164)	VC̥C̥	(0.442)
	/webbɯ/ ‘web’	VCC	(0.521)	VC̥C̥	(0.554)	VC̥C̥	(0.413)	VC̥C̥	(0.477)
	/bobbɯ/ ‘bob’	VCC	(0.636)	VC̥C̥	(0.320)	VC̥C̥	(0.667)	VC̥C̥	(0.921)
/d/	/kiddo/ ‘kid’	VCC	(0.387)	VC̥C̥	(1.251)	VCC	(0.937)	VCC	(1.699)
	/ɯddo/ ‘wood’	VCC	(0.457)	VC̥C̥	(1.161)	VCC	(0.983)	VC̥C̥	(1.411)
	/deddo/ ‘dead’	VCC	(0.368)	VNC	(1.043)	VCC	(0.894)	VNC	(0.125)
	/goddo/ ‘god’	VCC	(0.340)	VNC	(1.610)	VCC	(0.795)	VNC	(0.265)
/ʤ/	/badʤi/ ‘badge’	VCC	(0.574)	VC̥C̥	(0.204)	VC̥C̥	(0.405)	VC̥C̥	(1.071)
	/ɾidʤi/ ‘ridge’	VCC	(0.534)	VC̥C̥	(0.242)	VC̥C̥	(0.324)	VC̥C̥	(0.241)
	/edʤi/ ‘edge’	VCC	(0.493)	VC̥C̥	(0.272)	VC̥C̥	(0.390)	VNC	(1.791)
	/ɾodʤi/ ‘lodge’	VCC	(0.554)	VC̥C̥	(0.295)	VC̥C̥	(0.478)	VC̥C̥	(0.585)
/g/	/maggɯ/ ‘mug’	VCC	(0.470)	VC̥C̥	(0.403)	VC̥C̥	(0.553)	VC̥C̥	(0.722)
	/ʤiggɯ/ ‘jig’	VCC	(0.416)	VC̥C̥	(1.013)	VCC	(1.228)	VC̥C̥	(1.067)
	/hoggɯ/ ‘hog’	VCC	(0.461)	VC̥C̥	(1.008)	VC̥C̥	(1.125)	VC̥C̥	(1.476)

of the lexicon. Network 1, which was trained on Yamato, Sino-Japanese, and Foreign items, shows essentially the modern pattern, in which /dd,dʤ,gg/ $\succ$ /bb/ in loanwords. While the results from Network 2, which was trained on Yamato and Sino-Japanese items only, are rather inconclusive, the results from Networks 3 and 4 suggest that some combination of the following two factors had a role to play in establishing voiced geminates as being a valid phonological sequence in loanwords: early loans like /beddo/ that are variably attested with voiced geminates, allowing Japanese speakers to generalize the pattern to new loanwords; and the past tense forms of verbs, many of which contain a geminate /t/ and thus have the effect of increasing the frequency of words with /tt/ being presented to the network. All of the networks produced either VCC, VC̥C̥, or (in Networks 2 and 4) VNC adaptations for the test loanwords. While VNC is unattested as an established adaptation for voiced geminates, it can be seen as arising from the simultaneous satisfication of the constraint against voiced geminates, and the preservation of the input value of the [voi] feature in μ_2 of each network. In terms of OT constraints, Networks 2 and 4 have learned the ranking *DD, IDENT-SB(voi) $\gg$ IDENT-SB(nas), resulting in the geminate being changed into a nasal-stop cluster by these networks. Presumably, historical Japanese speakers never produced a VNC adaptation because the faithfulness constraint IDENT-SB(nas) was more highly ranked for them than it is in Networks 2 and 4.

While Network 2 does not show the expected differences in adaptation pattern by the place of articulation of the geminate stop (except for VNC adaptations for two of the words with /dd/), Networks 3 and 4 both adapt /bb/ and /dʤ/ as VC̥C̥, while /dd/ is adapted as either VCC (in Network 3) or VCC/VNC (in Network 4). Network 3, at least, has thus learned that /dd/ is the most acceptable out of the four possible voiced geminate obstruents, which accords with the historical attestation data. Since the training set for Network 3 contained only one example with a voiced geminate, namely /beddo/, the network must have derived the relative acceptability of /dd/ by generalizing not only from the presence of /beddo/, but also from the phonotactic patterns present in the Yamato and Sino-Japanese items it was trained on. Specifically, the network seems to be inferring, based on the greater number of training examples with /tt/ as compared to those with /pp/ or /ttʃ/, that geminate /dd/ should be more acceptable than /bb/ or /ddʒ/. What is interesting here is that Japanese speakers historically have found /ddʒ/ to be relatively acceptable as well, given that it begins to be attested in loanwords at about the same time as /dd/. This may be because in Japanese phonology, in voicing alternations like *rendaku* /dʒ/ is the voiced counterpart of /ʃ/ as well as /tʃ/, but the network cannot easily learn this fact since it is not trained on any examples of voicing alternations, and the featural representations of /ʃ/ and /dʒ/ differ by more than the value of the feature [voi]. Since /ʃʃ/ has a type frequency comparable to that of /kk/ (Figures 4.14 and 4.15), it may be that historical Japanese speakers were then able to infer that /ddʒ/ should be as acceptable as /dd/ or /gg/.

4.6 Conclusion

In this chapter I have examined how to model the process of loanword adaptation, focusing on the causes of frequency and similarity effects which have influenced adaptation patterns over time. While these kinds of effects can't be reproduced using the Core-Periphery or Cophonology models, because of the coarse-grained nature of lexical strata, they can be in connectionist models, in which strata are emergent collections of smaller-grained phonological neighborhoods. I then presented two connectionist models of a subset of the Japanese lexicon, the first of which demonstrates the nature of lexical representations in a CN, and the second of which explains the differences in adaptation patterns seen in loanwords with voiced geminates as being related to the higher type frequency of words in the Yamato and Sino-Japanese strata with geminate /tt/ and /kk/ as compared to /pp/.

In the next chapter, I shift focus from modeling individual speaker competence to modeling the speech community as a whole, as I consider the effects that the transmission of loans from speaker to speaker throughout a speech community has on the resulting nativizations that take place. After reviewing what is currently known about the large-scale structure of social networks, I then derive a simple probabilistic model for the expected global rate of nativization given the probability that an individual speaker will nativize a loanword. I find that transmission in general causes a nonlinear amplification of the effects of nativization at the level of the individual speaker. I then apply this model to the historical data from Chapter 2, showing that much of the variation in attested nativization rates over time for various non-native phonotactic patterns can be attributed to the effects of transmission.

CHAPTER 5

LOANWORD TRANSMISSION IN SOCIAL NETWORKS

In the last two chapters, I examined various models of loanword adaptation, arguing that any model which can plausibly be used to explain historical adaptation patterns must be one in which the production and perception of loanword tokens can be influenced by the existence of phonologically similar entries in each individual speaker's lexicon. However, I left unspecified what role transmission plays in the process of nativization. In this chapter, I will first examine different approaches to characterizing the structure of social networks. I will then develop a formal framework for examining the effects of transmission on the historical development of nativization patterns, showing how the nativization patterns discussed in Chapter 2 can be seen as arising from the cumulative effect of imperfect transmission of loanword variants through a network of speakers. I will show that the main effect of transmission is to increase the overall effect of nativization at the level of a single speaker, depending on the structure of the social network. This can result in the network as a whole adopting the nativized form of a loanword, even if the tendency for nativization at the level of the individual speaker is relatively small.

I will then look at how to explain the attested rates of nativization from the historical data in Chapter 2 using the transmission model. My method here is to reconstruct the rate of nativization at the individual speaker level, based on the rate of nativization in attested loanword tokens, and the expected amplification effect from the transmission model. It turns out that, before about 1890, the nativization of coronals before /i/ by individual speakers seems to have taken place at a qualitatively greater rate than the nativization of other non-native phonotactic patterns, such as palatal-/e/ sequences or voiced geminates; after this date, however, the rate of nativization was comparable to that of other phonotactic sequences. I will suggest some possible explanations for why this would be the case, focusing on the question of whether SP-level or CS-level processes (or both) were involved in the nativization of coronal-/i/ sequences. I will propose that the palatalization of coronals involved both SP-level and CS-level processes before 1890, but due to a large number of loanwords entering the language just before this time, as well as increased English education, after 1890 palatalization became an SP-level process only.

5.1 Characterizing social structure

There have been two main research programs looking at how social relationships are structured in a given community. The older strand of research has been done mainly by sociologists and has spawned a subfield of sociology known as *social network analysis* (Wasserman & Faust 1994, Scott 2000, Carrington, Scott & Wasserman 2005). Social network analysts view the various social relationships that people form within a community as constituting a *network* or *graph* (Harary 1972), where the individuals in the community correspond to the nodes of the network, and the relationships between them correspond to edges joining the individual nodes in the network (Figure 5.1). In modeling social networks as a graph, there are many possibilities for representing the properties of the social relationships between the members of a community. For example, the different kinds of relationships, such as acquaintance/friendship, parent-child, teacher-student, and so on, can be represented as different sets of links between the nodes of the graph, while the relative strength of the relationship, in terms of the amount of time and energy invested in maintaining the relationship, can be represented by associating a weight with each link, so that more highly-weighted links correspond to stronger ties (Granovetter 1973). The nodes themselves can likewise be characterized as either

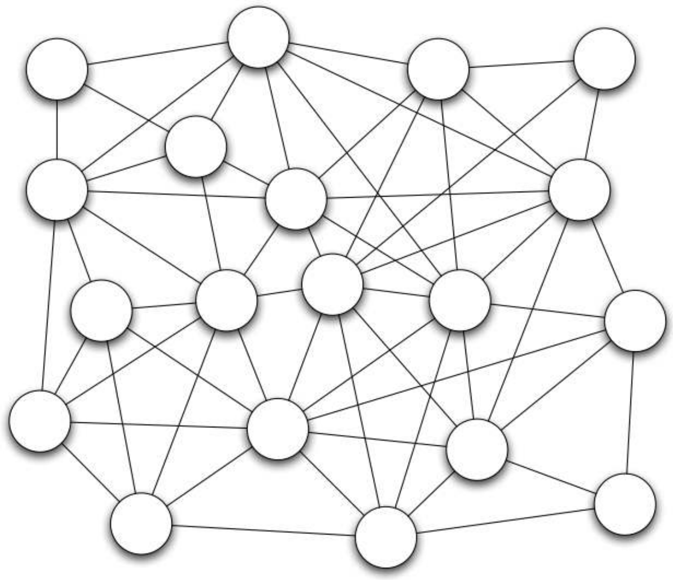


Figure 5.1: A hypothetical social network. Nodes represent speakers in the network, while edges represent social relationships between speakers through which communication can take place.

central or *peripheral*, depending on the number of connections to other nodes in the network. With this formalism in place, questions about how information is spread through the network can be made more precise, such as how long does it take for a social innovation to spread from a small group of initial adopters to the rest of the network, and how this is affected by the centrality of the nodes spreading the innovation (Valente 1995).

The other, and somewhat newer, research tradition also models social structure in terms of a graph, but instead of examining the properties of individual nodes in the network, the focus is instead on large-scale structural properties of the graph itself, in particular the *scale-free* nature of the degree distribution of many networks found in the real world (meaning that there are some nodes which are highly connected to many other nodes, and that there is no typical value for the number of nodes that a given node is expected to be connected to). These researchers, who are for the most part physicists and mathematicians, also look at other types of networks besides social ones, such as the structure of the World Wide Web, neural networks found in organisms like the nematode *C. Elegans*, and the power transmission grid in the United States, to give three examples cited in Newman (2003). While the majority of this research has been done in the past ten years, spurred by the publication of Watts & Strogatz (1998) and Albert & Barabási (1999), past researchers such as Yule (1925), Zipf (1935), and Simon (1955) have also looked at the properties of scale-free distributions (although not necessarily in the context of social or other kinds of networks).

5.1.1 Social network analysis in sociolinguistics

While mainstream researchers in sociolinguistics have preferred to talk about linguistic variation as being contingent on high-level social constructs such as class and gender (Labov 1972), there have been a few who instead analyze variation in terms of speakers' positions within their respective social networks. An early expression of this viewpoint can be found in Bloomfield (1933), who proposed the following thought experiment:

> Imagine a huge chart with a dot for every speaker in the community, and imagine that every time any speaker uttered a sentence, an arrow were drawn into the chart pointing from his dot to the dot representing each one of his hearers. At the end of a given period of time, say seventy years, this chart would show us the density of communication within the community. Some speakers would turn out to have been in close communication: there would be many arrows from one to the other, and there would be many series of arrows connecting them by way of one, two, or three intermediate speakers. At the other extreme there would be widely separated speakers who had never heard each other speak and were connected only by long chains

> of arrows through many intermediate speakers. If we wanted to explain the likeness and unlikeness between various speakers in the community, or, what comes to the same thing, to predict the degree of likeness for any two given speakers, our first step would be to count and evaluate the arrows and series of arrows connecting their dots.... We believe that the differences in density of communication within a speech-community are not only personal and individual, but that the community is divided into various systems of sub-groups such that the persons within a sub-group speak much more to each other than to persons outside their sub-group. Viewing the system of arrows as a network, we may say that these sub-groups are separated by *lines of weakness* in this net of oral communication. The lines of weakness and, accordingly, the differences of speech within a speech community are *local* due to mere geographic separation—and *non-local*, or as we usually say, *social*. (pp. 46-47, emphasis in original)

Of course, actual research studies using a network methodology necessarily fall far short of this idealized goal of recording every single speech act between all members of a speech community. The two most well-known examples of social network theory being applied to linguistic research, Milroy's (1987) study of sound change in Belfast English, and Eckert's (2000) study of Northern Cities Shift among Detroit high-schoolers, use various methods to approximate the actual structure of the network of communicative acts between the members of the speech communities being examined. Milroy (1987) defines a *network strength scale* which expresses the degree to which a particular speaker is more or less involved with the social networks in their local neighborhood. This scale consists of five indicators, each of which is scored as either 0 or 1 (Milroy 2002: 555):

1. Membership of a high density, territorially based group (e.g. a bingo or card-playing group, a gang, or a football team or football supporters' club);
2. Having kinship ties with more than two households in the neighborhood;
3. Same workplace as at least two others from the neighborhood;
4. Same workplace as at least two others of the same gender from the neighborhood;
5. Voluntary association with workmates in leisure hours.

Low-scoring individuals have few social connections with other neighborhood residents, while high-scoring individuals have multiple work and family connections with other members of the community. Milroy then found that this network score was positively correlated with various linguistic variables among Belfast speakers, such as (a) (the degree of backing in the production of /a/) and (th) (percentage deletion of intervocalic /ð/). She concludes that the degree to which an individual is integrated into their local social networks has a direct influence on their language behavior, since the more closely connected an individual is with the other members of their neighborhood, the more likely they are to favor vernacular variants in their own speech.

However, Murray (1993) points out several flaws in Milroy's (1987) methodology and analysis, although the most serious claims of irregularities in Milroy's statistical analyses were later retracted (Butters 1995). One of the remaining issues is that Milroy's network scale is arbitrarily defined and not an interval scale (meaning that the difference in social network participation between a score of 1 and 2 is not necessarily the same as the difference between a score of 2 and 3, or 3 and 4, etc.). This makes it difficult to interpret any correlations between these network scores and linguistic variables. In addition, Murray argues that the network scale collapses together two very different kinds of individuals, since low scorers could either be individuals who are isolated from any social relationships at all, or they could be upwardly mobile and breaking their ties with the local neighborhood in order to move into the middle class. It seems implausible that both types of speakers would have similar rates of usage of vernacular variants in their own speech. Murray then presents a reanalysis of Milroy's data, and finds that a better predictor of vernacular usage among Belfast speakers is the sex of the speaker, with men tending to use the vernacular more than women. Milroy (2002) briefly responds to these criticisms by stating that Murray's beliefs about how statistical analysis should be conducted are not uncontroversial, but does not go into any detail on this point.

Eckert (2000), in her study of Northern Cities Shift (NCS) among suburban Detroit high-school students, provides a more direct approach to applying social network analysis to studying language change. Her method is to do an ethnographic study of the social categories which the students divide themselves into ("jocks", "burnouts", and "in-betweens"), as well as constructing a sociogram from students' reported friendship ties with other students. Eckert finds that the students' self-reported classifications of themselves and others corresponded with their relative position in the sociogram, with the in-betweens forming two clusters of students connecting the clusters of jocks and burnouts. The students' positions in the social network also correlated with the degree to which they exhibit NCS characteristics in their own speech, with the burnouts' vowel systems being the most advanced with regards to NCS, and the jocks' vowel systems being the least shifted.

To examine loanword transmission using a social network approach, the attested patterns of nativization seen in a particular loanword would be accounted for by constructing a graph representing the social relationships among a selected group of historical speakers at the time of the loanword's first attestation (or, if we're feeling ambitious, all speakers of Japanese at that date), and then examining how the loanword spread from speaker to speaker through the network. It is, of course, completely infeasible to model social networks to this level of detail with historical data, given that the needed information about the relevant relationships between speakers rarely, if ever, survives to the present day.[1] Besides, even if there were enough information to reconstruct the actual network structure of, say, a group of early Meiji period college students, documenting the spread of a loanword through this particular network would not necessarily allow us to generalize to similar processes occurring in other social networks. It may be that a particular feature of the structure of that network, say, a highly clustered clique centered around one of the initial borrowers, is what led to the established form of the loanword being nativized or not.

Thus what is needed is a neutral model of the structure of a typical social network, which would match the large-scale structural properties of real-world networks. Then by analyzing the expected behavior of a loanword as it is transmitted through such a network, along with computer simulations of the spread of a loanword using many different randomly-chosen networks with the same structural parameters, it would be possible to characterize more precisely the effect that the transmission of the loanword through the network has on the resulting process of nativization. In order to develop such a model of network structure, it is necessary to examine the large-scale properties of such networks affecting the transmission of information from one node to another, which I will now turn to in the next section.

[1] The study of syntactic change in Middle English by Bergs (2005) is a rare exception. Of course, modern-day studies of currently propagating loanwords could potentially collect this type of data as well, using methods similar to those of Eckert (2000).

Table 5.1: Typical values of graph parameters for random and social networks. N = number of nodes; m = number of edges; M = number of possible edges = $\binom{N}{2}$; $\langle k \rangle$ = expected node degree = $\frac{2m}{N}$.

Parameter		Random graphs	Social networks
$\frac{m}{M}$	density	low ($\ll 1$)	low
ℓ	average path length	low ($\approx \frac{\log N}{\log \langle k \rangle}$)	low
C	clustering coefficient	low ($\approx \frac{\langle k \rangle}{N}$)	high ($\gg \frac{\langle k \rangle}{N}$)
p_k	degree distribution	Poisson	exponential, power law (?)

5.1.2 The large-scale structure of social networks

Recent research in the physics literature on the statistical mechanics of real-world social networks has found that these networks[2] have similar large-scale structural properties (Watts & Strogatz 1998, Albert & Barabási 1999). This structure can be characterized using a number of different parameters, listed in Table 5.1. Here I have compared the properties of real-world social networks with those of the *random graph* model of Erdős & Rényi (1960), which has been used in the past as a simple model of network structure (Albert & Barabási 1999). A random graph is generated by creating a set of n nodes, and then including each possible edge between each pair of nodes with probability p (which Erdős & Rényi call the $G(n,p)$ model).[3] With a relatively small value of p,[4] random graphs can replicate the first two properties of social networks listed in Table 5.1. The *density* of the graph, that is the number of edges divided by the number of possible edges, is relatively low, while the *average path length*, the average number of edges between two nodes picked at random, scales with $\log n$ (meaning that, as the number of

[2] As well as other kinds of networks found in the natural world, such as the power distribution network, or gene expression networks, among others (Amaral, Scala, Barthelemy & Stanley 2000).

[3] A very similar model is the $G(n,M)$ model, where a graph is chosen at random from the set of all graphs with n nodes and M edges.

[4] Erdős & Rényi (1960) found that $p = \frac{\log N}{N}$ is a threshold value with regards to whether or not a random $G(n,p)$ graph is connected. If $p < \frac{\log N}{N}$, then most $G(n,p)$ graphs will be disconnected, whereas if $p > \frac{\log N}{N}$, then the graphs will typically have a single giant component which contains most of the vertices in the graph.

nodes in a graph increases, the average path length grows much more slowly). These two properties in combination are known as the *small-world effect* (Watts & Strogatz 1998), named after the famous "small world" experiment conducted by social psychologist Stanley Milgram (Milgram 1967, Travers & Milgram 1969). However, despite being rather counter-intuitive on the face of it, the small-world effect is actually a typical property of random graphs (Newman 2003).

Yet social networks turn out to differ from random graphs in two important ways. First, real-world networks have a high *clustering coefficient* (Kossinets & Watts 2006), which measures how well connected (or "cliquish") the neighborhoods of each node are (where the *neighborhood* of a node n is the set of nodes that are directly connected to n). This is defined as the number of edges between the nodes in the neighborhood divided by the number of possible links:

$$C_i = \frac{2|\{e_{jk}\}|}{k_i(k_i - 1)} : n_j, n_k \in N_i, e_{jk} \in E \tag{5.1}$$

(where E is the set of edges in the graph, $N_i = \{n_j : e_{ij} \in E\}$ is the set of neighbors of node n_i, and $k_i = |N_i|$ is the degree of node n_i). While random graphs tend to have a very low clustering coefficient ($C_{random} \approx \frac{\langle k \rangle}{N}$), social networks tend to have a much higher clustering coefficient ($C \gg C_{random}$; Watts & Strogatz 1998). What this means is that in social networks, two nodes are much more likely to be connected to each other if they are both connected to some common third node, whereas in random graphs, the likelihood that two nodes are connected does not depend on whether there are any neighbor nodes in common between the two. More informally, social networks tend to be made up of clusters of smaller highly-connected cliques, while random graphs are far less likely to exhibit this kind of structure.

Second, social networks typically have a different *degree distribution* than random graphs. The degree of a node is the number of nodes it is connected to in the graph, while the degree distribution is a probability distribution over the degrees of all of the nodes in a graph. In social networks, the degree distribution is typically an exponential or power-law distribution (Albert & Barabási 1999), while a random graph has a Poisson degree distribution. What this means is that in a random graph, there is a typical or "average" degree value, and the degree of most nodes will cluster around this value, while in a real-world network, node degrees vary over a large range of values, with most nodes having a relatively small degree and a few nodes having a relatively large degree, and there is no typical degree value. Figure 5.2 compares the cumulative degree distributions of $G(N,p)$ graphs with Barabási & Albert (1999) scale-free graphs, for randomly-generated graphs of each type with $N = 500$ nodes. It can be seen in this plot that, for random graphs, the probability of finding a node with a given degree decreases rapidly with increasing degree, such that it is extremely unlikely to find a node having a degree greater than about 15-20 for the randomly-generated networks shown here. For scale-free graphs, on the other hand, this probability decreases much more slowly, and it is possible to find nodes in such networks with much higher degrees of 40 or even 50. These "hub" nodes connecting a significant fraction of the total number of nodes in the network, while common in scale-free graphs, practically never occur in random graphs.

It is debatable, however, to what degree the types of social networks relevant in spreading loanwords are themselves scale-free (having a power-law distribution). Whether a particular network is scale-free or not seems to depend on whether there are constraints on creating and maintaining links in the network (Amaral et al. 2000). If there is no cost in creating a new link or maintaining an existing link, then the resulting network will be scale-free, as seen in the citation network of scientific research (Redner 1998) and the network of human sexual contacts (Liljeros, Edling, Amaral, Stanley & Åberg 2001). However, if there is a cost involved, then the nodes of the network will have an exponential or Gaussian distribution of connectivities instead, which

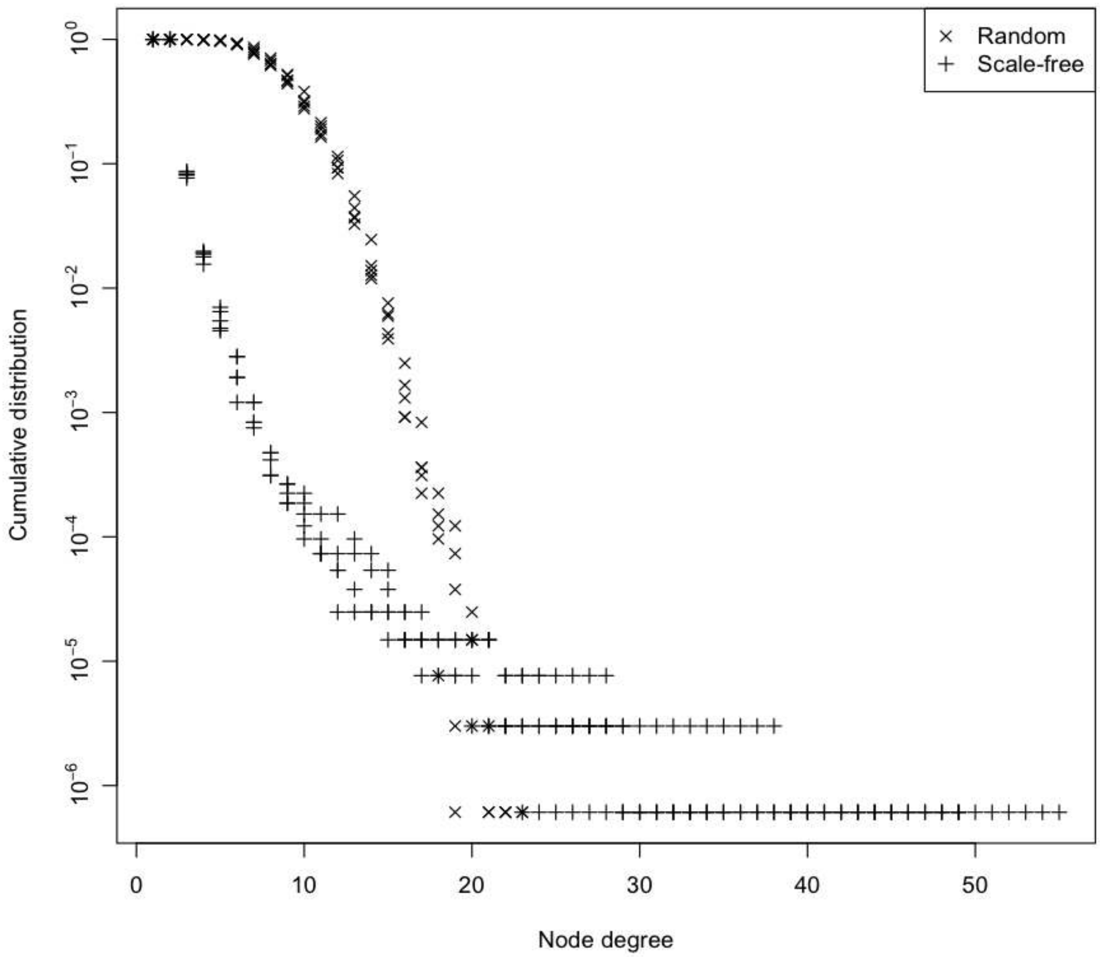


Figure 5.2: Log-linear plot of cumulative degree distributions of random and scale-free graphs. Five graphs of each type were randomly generated with $N = 500$. For $G(N, p)$ graphs, $p = 1.5\frac{\log N}{N}$.

is what is seen in friendship and acquaintance networks, where presumably the cost involved is the amount of time needed to maintain a friendship. The question is to what degree loanwords spread through low-maintenance contacts (like chance encounters with strangers, or through mass media) vs. high-maintenance contacts (like work relationships or friendships), since this will determine whether the scale-free quality of low-maintenance contact networks is critical to the establishment of adaptation patterns.

5.2 A formal framework for modeling loanword borrowing

In order to look at the effects of loanword transmission, it is necessary to model not only the adaptation of a new loanword by a single speaker, but also the dynamics of loanword propagation among a group of such speakers embedded within a social network. Given the discussion above, such a network can be represented using a graph consisting of a set of nodes V and a set of edges E between them, where e_{ij} denotes an edge between nodes v_i and v_j. Each node can be taken to represent a single speaker in the social network, while each edge represents a social tie (friendship, work relationship, etc.) through which communication between speakers takes place. To simplify things, I will assume that the graphs representing social networks are *simple*, meaning that each edge always connects two distinct nodes, and for any pair of nodes there is at most one edge between them;[5] and *fully connected*, meaning that for any pair of nodes in the graph, there is at least one set of edges forming a path between them.[6]

[5]This rules out multiple edges between a given pair of nodes, as well as an edge having the same node for both endpoints, forming a loop.

[6]Granovetter (1973) and subsequent work in sociology has explored more complex models where it is possible for two nodes to have more than one link between them (which can represent multiple relationships between nodes, such as a friendship and work relationship), as well as models where links have a numerical weight associated with them representing the strength of the link (so that links with higher weights are more important to the two participants than links with lower weights). Granovetter argues that weak ties are more likely to connect members of different social groups, and thus are important for the diffusion of innovations between these groups.

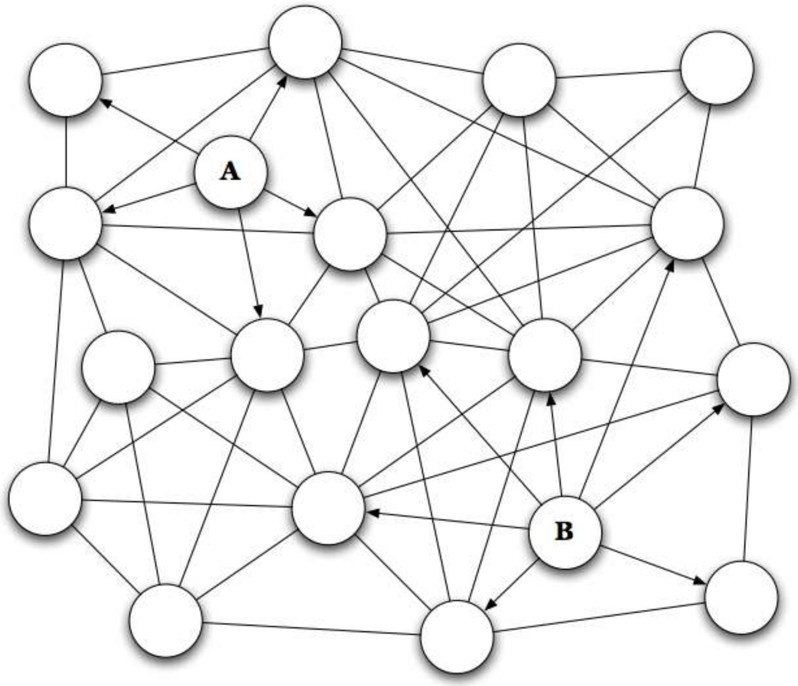


Figure 5.3: Loanword transmission in a graph representing a social network. Speakers A and B introduce a new loanword, which can then spread to other individuals they are in contact with (as indicated by the arrows). These speakers can then spread the loanword to their neighbors in the graph, and so on.

The borrowing of a new loanword can be viewed as the introduction of the word at one or more of the nodes of the graph (representing the adaptation of the word by those speakers), followed by the gradual spread of the word to the rest of the nodes along the edges of the graph (representing the transmission of the word from speaker to speaker), as shown in Figure 5.3. It turns out that the effect of transmission on the resulting nativization of a loanword can vary greatly, depending on the structure of the graph.[7] In the discussion that follows I will first consider the simple case of a *line* graph, deriving a prediction for the expected rate of nativization in the entire network given the probability that an individual speaker nativizes on the perception or production of an individual loanword token. I will then extend this probabilistic model to random *trees*,

[7]Nowak (2006) finds a similar result in his analyses of evolutionary processes taking place on different types of graph structures.

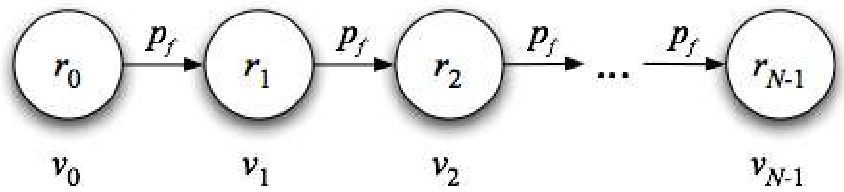


Figure 5.4: A line graph with N speakers. Each speaker v_i chooses an F-token target for production with probability r_i, then with probability p_f that target is perceived as an F-token by speaker v_{i+1}. In the process of learning the loanword, v_i estimates its value of r_i from the ratio of F-tokens to total tokens produced by v_{i-1}.

and then to more complex random graphs having similar structural properties as the real-world networks discussed in section 5.1.2.

5.2.1 Transmission in line graphs

I will first consider a very simple case of a graph, namely a *line* graph. This consists of N speakers arranged in a line, with the initial borrower at one end of the line, and with each speaker communicating only to the node to the right of them in the line (Figure 5.4). This graph is very similar to the one from the game of “Telephone” considered in section 2.2 of Chapter 2, only instead of each speaker always producing either the nativized (N) or non-nativized (F) form of the loanword, instead each speaker has an associated probability r_i that they will choose the F form as a target for production (so that an r_i of 0 would mean that that speaker would always target the N form, an r_i of 0.5 would mean that the speaker would choose both the N and F production targets equally frequently, and so on). This probability represents each speaker’s knowledge of both the phonological form of the loanword, and of the level of variation of other speakers’ productions of the loanword. I am using probabilities here as a way of simplifying the problem of how to model the confluence of all of the different linguistic and cognitive factors which determine how a particular speaker produces a given loanword at a particular time. I am emphatically *not* suggesting that these factors are inherently random or unknowable; rather, I am abstracting away from the details of any particular model of loanword adaptation and lexical representation, in order to focus on the effects of transmission only.

I will also assume that, for every utterance that is produced, there is another probability p_f that, if the utterance target intended by the speaker was the F form of the loanword, then it will be successfully perceived as such by the listener.[8] This is meant to represent the effects of phonetic biases and phonological processes taking place during the perception and production of non-native segments and phonotactic sequences that tend to cause nativization of these elements in a loanword. If the target was the N form of the loanword, then I'm assuming that it will always be perceived as such by the listener, since the N form by definition contains only licit segments and phonotactic sequences, and so presumably would not pose the same kinds of problems for production and perception that non-native segments and phonotactic sequences would. (Note that I'm restricting my attention here to loanword variants which can unambiguously be classified as either N or F, like /tiːmɯ/ (F) and /ʧiːmɯ/ (N) for *team*, and for which there is at most one foreign element in their phonological representations.) A p_f of 0 would mean that every utterance is always perceived as the N form by the listener, even if the speaker intended to produce an F token, while a p_f of 1 would mean that every F token is successfully perceived as such by the listener. We can think of p_f as really being composed of two different probabilities: the probability that the speaker

[8]It might seem more intuitive at first to instead talk in terms of the probability $p_n = 1 - p_f$ that an F form will be nativized (turned into an N form) during each act of transmission. But it turns out that expressing equations 5.4 and 5.6 in terms of p_n makes the algebra involved rather unwieldy. This is because there are two ways in this model for a listener to be exposed to an N form: when a speaker chooses an N target for production, and when a speaker chooses an F target which then gets nativized. The probability of an N token being produced by a speaker v_i is then $(1 - r_i) + r_i p_n$. However, there is only one way for a listener to be exposed to an F form, namely, when a speaker chooses an F target, and this target is correctly produced and perceived as an F token. The probability that an F token will be produced by v_i is then simply $r_i p_f$.

will accidentally nativize during production (p_{produc}), for example by missing a gestural target and producing something more like [ʧi] than [ti]; and the probability that a successfully-produced non-native token will be misperceived as the nativized variant by a listener (p_{percep}), for example by misinterpreting the aspiration noise following a [t] as frication noise and thus interpreting the [t] as a token of /ʧ/. Then p_f as it is defined here would be some function of p_{produc} and p_{percep} (such as $p_f = p_{produc} \times p_{percep}$). However, to simplify the analysis of this model, I will use the single parameter p_f only to cover the cases of both misproduction and misperception.

Notice that when $p_f = 0$, every speaker (excluding the initial borrowers) will be exposed only to the N form of the loanword. This case would correspond to the implicit assumption often made in loanword research that nativizations are always made by a single speaker (either the initial borrower, or an L1 speaker in contact with an L2 speaker or L1/L2 bilingual). As I discuss in section 5.3, this would predict—contrary to the nativization data from Chapter 2—that only the N form of the loanword should ever be attested, and there should be no variation between the N and F forms.

It is also important to note that while the value of r_i for each node v_i is specific to that node, and is learned based on its interactions with nearby nodes in the network, the parameter p_f is instead a global parameter which is applied to every utterance produced in the network. This is obviously a gross simplification of what happens in real-life speech communities, in which different speakers will have very different abilities to perceive and produce non-native segments and segment sequences. For example, a very conservative speaker might prefer to nativize all such foreign phonemic structures, whereas a more innovating speaker might prefer to produce all of these elements as closely as possible to their L2 source. In fact, Bloch (1950), in his phonemic analysis of modern Japanese, states his generalizations in terms of two types of speakers, a “conservative” speaker who always nativizes loanwords, and an "innovating" speaker who never nativizes.[9] As well, Akamatsu (1997) notes that different speakers have different strategies for producing the sequences [ɸa,ɸi,ɸe,ɸo] (only [ɸɯ] occurs in native words, and in fact [ɸ] is an allophone of /h/ before /ɯ/ in the Yamato, Sino-Japanese, and Mimetic strata). Some speakers produce these sequences as [ɸa,ɸi,ɸe,ɸo], with [ɸ] being followed directly by the vowel, whereas other speakers insert an epenthetic [ɯ] instead, producing [ɸɯa,ɸɯi,ɸɯe,ɸɯo]. In principle, these types of speaker-specific tendencies to nativize or not could be represented by having an individual nativization probability p_i associated with each node v_i. This would make the model considerably more complex, however, since instead of the single parameter p_f, there would be N parameters $p_0, p_1, \ldots p_{N-1}$. To simplify the analysis of this model, then, I am setting these issues aside for the time being and simply using the global parameter p_f to represent the rate of nativization at the level of a single individual in the network.

Suppose now that each node, in the process of learning the loanword, estimates its value of r_i from the frequency of F tokens in the total set of tokens of the loanword that it is exposed to. For example, suppose the initial borrower v_0 borrows the word *team*, and produces 10 tokens of this word for speaker v_1, four of which have the phonetic form [ʧiːmɯ] and the other six having [tiːmɯ]. Then v_1 will have a value of 0.6 for r_1. Now suppose that the loanword is allowed to propagate from speaker to speaker until every speaker has learned the loanword, and we then want to know what is the behavior of this mini-speech community with respect to the nativization of the loanword: namely, if a

[9]Bloch (1950) did this because he believed that there was no legitimate basis from the viewpoint of the speaker for excluding loanwords from the phonemic analysis of the native phonological system of a language, as some other Structuralists argued for at the time (e.g. Fries & Pike 1949), since the typical speaker will not necessarily know the etymology of the words in his mental lexicon, and thus cannot be expected to reliably classify words by stratum. This skepticism is echoed in recent works such as Rice (1997) and Ota (2004). However, the results of the stratum classification network in Chapter 4 show that the etymological classifications of Japanese words are still reflected to a great degree in the phonological similarity structure present in the Japanese lexicon. This suggests that the learnability problem is not as dire as Bloch makes it out to be.

node is picked at random, how likely is it to produce an F token of the loanword? This will depend on the expected values of r_i at each node. Since the speakers are arranged in a line, with each speaker only hearing tokens of the loanword from the speaker before it in line, then it is clear that the value of $\langle r_i \rangle$,[10] for $i > 0$, depends only on the values of r_{i-1} and p_f. Supposing speaker v_{i-1} produces n tokens of the loanword (with n being sufficiently large to avoid discretization effects), then approximately nr_{i-1} of these will be F-token targets, and of these, $nr_{i-1}p_f$ will be correctly produced by v_{i-1} and then perceived as F tokens by v_i. Thus v_i can be expected to estimate r_i as

$$\langle r_i \rangle = r_{i-1} p_f. \tag{5.2}$$

Substituting for $i = 1,2,3,\ldots$, we find that $r_1 = r_o p_f$, $r_2 = r_1 p_f = r_o {p_f}^2$, $r_3 = r_2 p_f = r_o {p_f}^3$, and so on. More generally,

$$\langle r_i \rangle = r_o {p_f}^i. \tag{5.3}$$

If a node v_i is picked and made to produce a loanword token, it will produce an F token with probability $r_i p_f$. Thus to find the expected probability $\langle q_f \rangle$ that a randomly chosen node (assuming that each node is equally likely to be chosen) will produce an F token, we need to sum the values of $r_i p_f$ for all nodes v_i and divide by the number of nodes in the network:

$$\langle q_f \rangle = \frac{1}{N} \sum \langle r_i \rangle p_f = \frac{r_0 p_f}{N} \sum_{k=0}^{N-1} {p_f}^k. \tag{5.4}$$

Equation 5.4 has two important implications. First, note that $\langle q_f \rangle \leq r_0 p_f$ for all values of p_f and r_0 between 0 and 1. In other words, the expected global rate of nativization is never less than the rate of nativization at the level of an individual speaker, as represented by p_f, and in fact is in many cases greater than the individual rate of nativization. In this model, then, the transmission of a loanword through a social network can result

[10]The notation $\langle x \rangle$ denotes the expected (or mean) value of x.

in more nativization taking place than would be expected if only a single speaker were involved. Second, nativization at the level of the entire network is a nonlinear function of the individual rate of nativization. This is because equation 5.4 contains a term ${p_f}^k$, where k is the distance (in terms of the number of graph edges) for a given node from the initial borrower v_o. This term represents the cumulative effect of misperception and misproduction during the transmission of a loanword from speaker to speaker, and has a greater effect as the number of nodes in the network increases, as shown in Figure 5.5, which is a graph of equation 5.4 for various values of p_f and N (and with $r_o = 1$).[11] Here we can see that in the case $N = 1$ (i.e. the only speaker in the network is the initial borrower), $\langle q_f \rangle = p_f$; but as N increases, the curve for the value of $\langle q_f \rangle$ moves farther away from the diagonal line representing the $N = 1$ case. For $N = 10$, this curve remains relatively low for $p_f < 0.75$, then increases sharply, meaning that even if speakers produce and correctly perceive the F form of the loanword up to three-quarters of the time, it can still result in a randomly-chosen speaker from the network tending to produce the N form instead. Transmission in the line network thus can result in a strong amplification of the effects of nativization at the level of the individual speaker, and this effect increases the more speakers there are in the network.

To test the predictions of this model, I performed computer simulations of loanword transmission in line networks with various values of p_f. In all of these simulations, the nodes are in one of two states, w_0 (does not know the loanword) and w_1 (knows the loanword). At the start of each simulation, v_0 is the only node in state w_1, and $r_0 = 1$. All of the other nodes $v_1, v_2, \ldots v_{N-1}$ start off in state w_0. On each time step of the simulation, each node v_i in state w_1 produces a token of the loanword for its neighbor to the right in the network. v_i chooses an F form as a production target with probability r_i (and thus an N form with probability $1 - r_i$). While N forms are always correctly

[11] With $r_0 < 1$, the graph of equation 5.4 looks similar, but with the rightmost endpoint at $(1, 1)$ at which all of the curves converge shifted down to $(1, r_0)$.

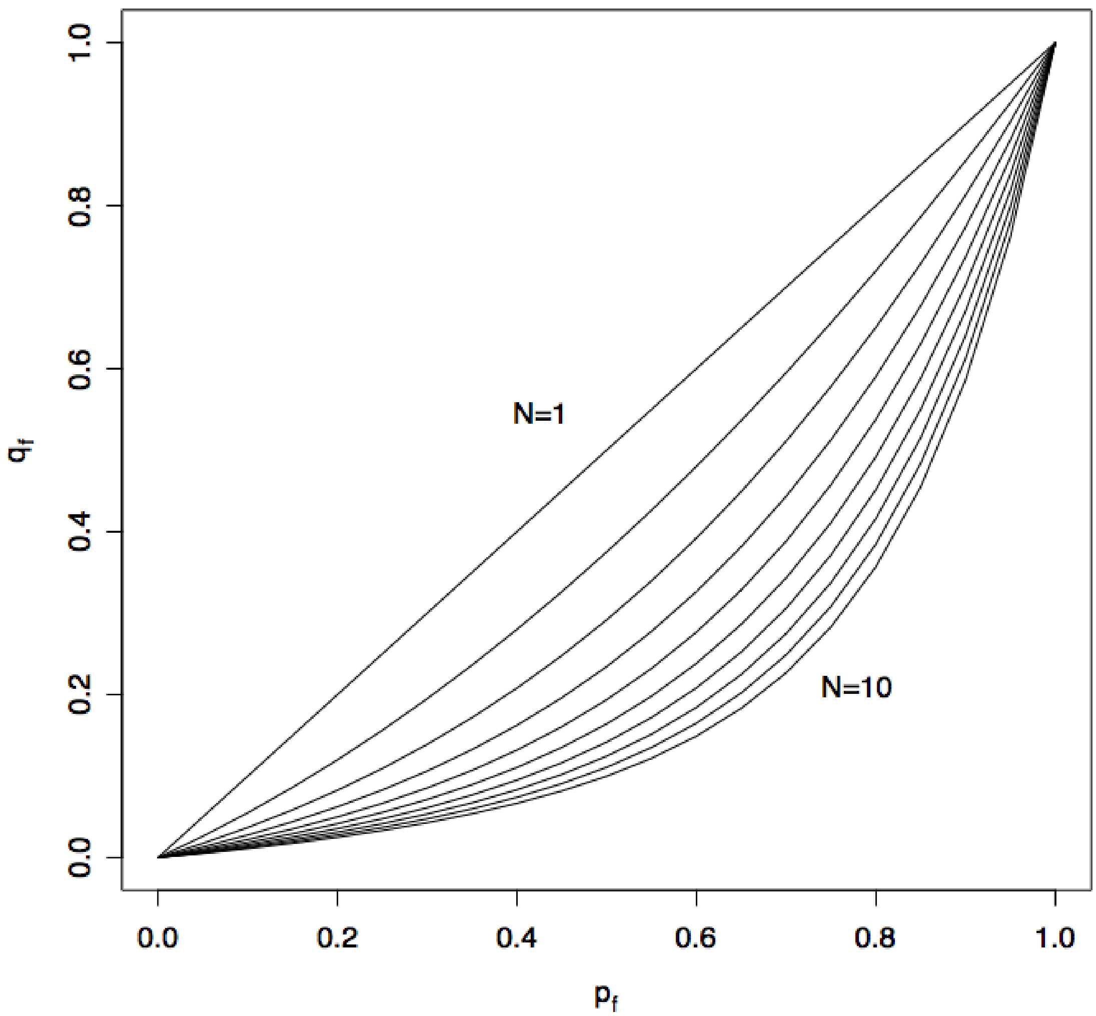


Figure 5.5: Predicted values of q_f for various values of p_f in a line network, with $N = 1, 2, \ldots 10$ and $r_o = 1$.

perceived by neighbor nodes, F forms are correctly perceived with probability p_f. Then, for each of these neighbor nodes v_j, if they are in state w_0, there is a probability p_{learn} that they will switch to state w_1. At this time, they estimate their value of r_j based on the ratio of F tokens to total tokens that they have perceived up to this point, and starting on the next time step, they will begin producing tokens of the loanword as well. The simulation ends when all nodes are in state w_1 (in other words, when all nodes have learned the loanword).

Figure 5.6 shows the results of these simulations. Here, for each value of $p_f = 0, 0.05, 0.1, \ldots 1$, 20 line networks with 10 nodes each were constructed, and then loanword transmission was simulated as described above. At the end of each simulation, the mean value of r_i for all nodes was recorded, and then this was multiplied by the value of p_f for that simulation to find the probability q_f that an F-token will be produced by a randomly selected node in that network. As the graph shows, the predicted value of $\langle q_f \rangle$ from equation 5.4 corresponds well with the actual results from simulations when $p_f \leq 0.5$ (and also for the trivial case of $p_f = 1$, where all nodes learn to produce F-tokens only). For $0.5 < p_f < 1$, there is a great deal more variability in the resulting value of q_f. For example, the results from the simulations for $p_f = 0.9$ range from less than 0.2 to more than 0.8. This variation arises to some extent from the structure of the line network. Since each node is connected to at most two other nodes (the left neighbor and the right neighbor), the nodes closer to v_0 have a greater effect on the value of q_f than nodes farther away. For example, if the node v_1 by chance happens to have heard only N-tokens of the loanword from v_0 when it learns the loanword, then it will estimate its r-value as $r_1 = 0$, and will only produce N-tokens for its neighbor v_2 (which will then only produce N-tokens after it learns the loanword, having never heard any F-tokens from v_1). This will have the effect of making the actual value of q_f in this particular network be $q_f = \frac{r_0}{N}$, which will be much lower than the expected value $\langle q_f \rangle$ for a high

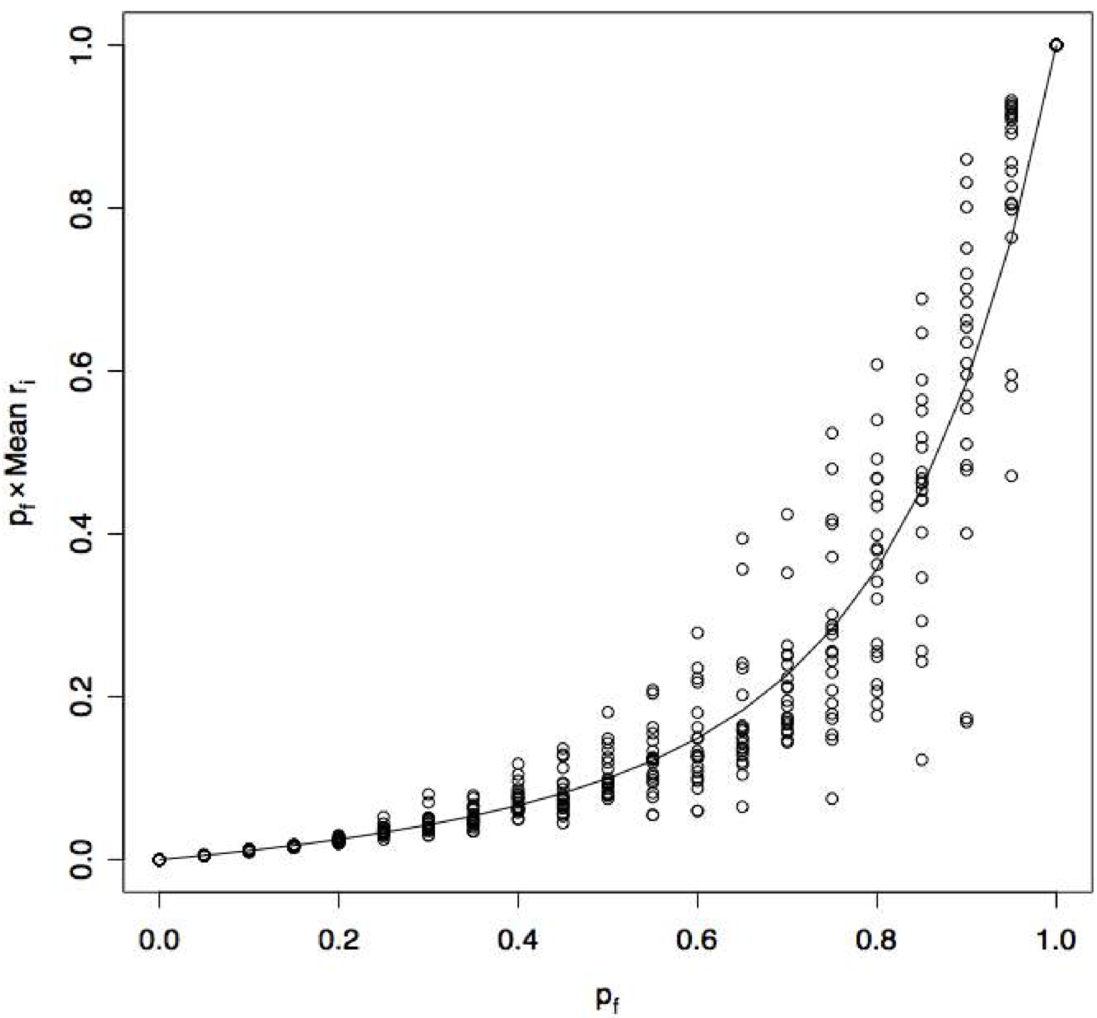


Figure 5.6: Results of computer simulations of loanword transmission in line graphs. $N = 10$, $r_0 = 1$, $p_{learn} = 0.05$. Each plotted point corresponds to a single randomly-generated graph; 20 graphs were simulated for each value of p_f. Line indicates predicted $\langle q_f \rangle$ values for a line graph with $N = 10$ and $r_o = 1$.

value of p_f. In the next section, it will be seen that more realistic social network models, such as random and small-world graphs, do not result in as much variability for the actual values of q_f.

5.2.2 Transmission in random graphs

The line graph discussed in the previous section is of course a poor model for the structure of a real-world social network, since speakers can generally have more than two neighbors, multiple paths between a given pair of speakers can occur, and so on. It is a simple matter to generalize the line model developed in the previous section to *trees*, which are graphs in which each pair of nodes is joined by a unique path (Harary 1972). Figure 5.7 shows an example of a tree, where the node labeled v_0 is the initial borrower. Since, for each node v_i in the tree, there is only one path leading back to v_0, v_i will be exposed to the loanword only from the immediately preceding node along the unique path from v_i to v_0 (that is, the node that is one link closer to v_0). Then, just as in the line network, the expected rate at which v_i will produce F tokens will depend only on the values of p_f and the distance (in terms of the number of links) from v_0 to v_i. By this reasoning, for all nodes of distance k from the initial borrower v_0, the expected rate of F-token production is

$$\langle r_k \rangle = r_{k-1} p_f = r_o {p_f}^k. \tag{5.5}$$

To find the expected value of q_f for a tree, let n_k be the number of nodes of distance k from the initial borrower v_0. For example, in the tree depicted in Figure 5.7, $n_0 = 1$, $n_1 = 4$, and $n_2 = 8$. Then there are n_0 nodes with an r-value of r_0, n_1 nodes with an r-value of r_1, and so on, and thus

$$\langle q_f \rangle = \frac{1}{N} \sum_{k=0}^{d} n_k r_k p_f = \frac{p_f r_o}{N} \sum_{k=0}^{d} n_k {p_f}^k, \tag{5.6}$$

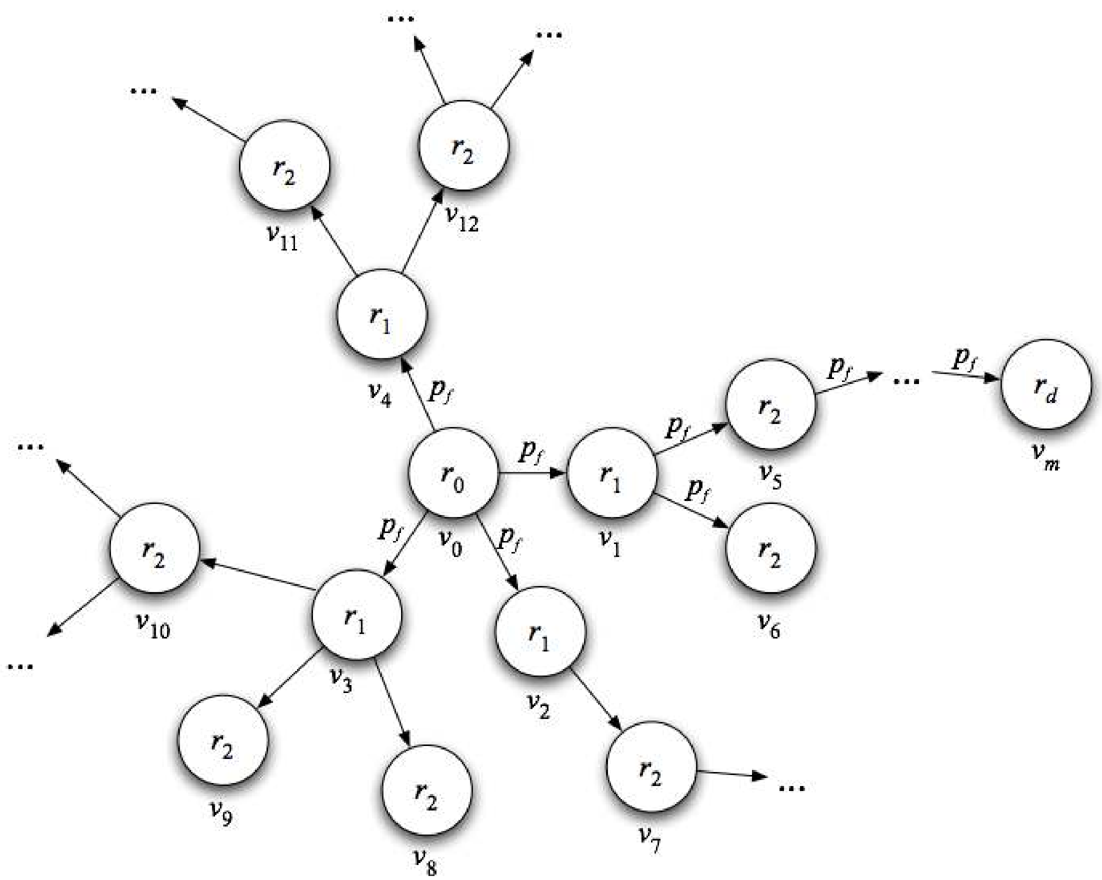


Figure 5.7: A tree of N nodes. The value of r for each node depends only on its distance from the initial borrower v_0, so that all nodes that are exactly one link away from v_0 have an r value of r_1, nodes that are two links away have r_2, and so on. d is the maximum distance from v_0 to any of the nodes in the tree.

where d is the maximum distance from v_0 in the tree. This equation is similar to (5.4), in that both are polynomials in p_f. This means that loanword propagation in trees will result in a similar nonlinear amplification of nativization as seen in line networks. To what degree nativization is increased will depend on the values of n_k and d: the more nodes there are at a farther distance from any of the initial borrowers, the more nativization will take place and the more nonlinear the $\langle q_f \rangle$ curve will become.

The tree model can be used to approximate the rate of nativization expected in any arbitrary graph. This is because, for any graph, there exists a *spanning tree*, which is a tree containing all of the nodes, and some subset of the edges, in the original graph (Gibbons 1985). In other words, it is possible to take a graph with a single initial borrower v_0, find a spanning tree for this graph rooted at v_0, and then count the number of nodes that are $1, 2, \ldots d$ links away from v_0 to find the values for $n_1, n_2, \ldots n_d$ in equation 5.6. A graph with multiple initial borrowers can similarly be converted into a set of trees, with each tree rooted at one of the initial borrowers (Figure 5.8). In fact, using the *breadth-first search* algorithm from Gibbons (1985: 35), this set of spanning trees can be constructed so that the minimum distances from each node to the closest initial borrower in the graph are preserved. In other words, if a node is of distance m from its closest initial borrower, then it will still be of distance m in the corresponding spanning tree. Then the values of $n_1, n_2, \ldots n_d$ can be found by counting the number of nodes in the original graph that are at least one link away from one of the initial borrowers, then the number of nodes that are at least two links away, and so on. Table 5.2 and Figure 5.9 shows the results of this process for $G(N, p)$ random graphs with 500 nodes and various numbers of initial borrowers. It can be seen that, for most networks of this size, the majority of nodes are either one or two links away from an initial borrower, with the average distance to an initial borrower decreasing as the number of borrowers increases.

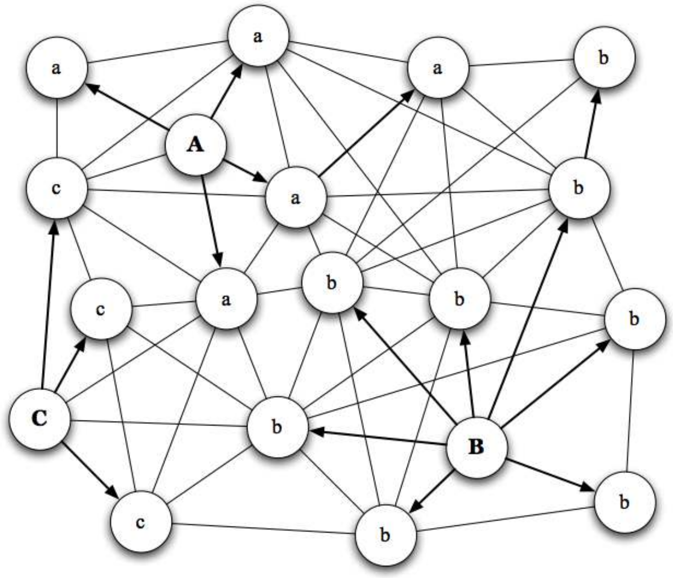


Figure 5.8: Splitting a graph into a set of trees rooted at the initial borrowers. The nodes labeled A, B, and C represent initial borrowers, while the arrows represent the links for one possible set of spanning trees rooted at these borrowers and containing the rest of the nodes in the network.

Table 5.2: Mean values of $n_1 \dots n_5$ in $G(N, p)$ random graphs for various values of n_0 and $N = 500$. Each entry represents the means from 100 randomly-generated networks.

n_0	Mean n_1	Mean n_2	Mean n_3	Mean n_4	Mean n_5
1	9.22	76.48	300.12	112.32	0.86
6	52.51	274.65	165.82	1.02	0.00
12	100.48	328.64	58.80	0.08	0.00
25	178.71	285.40	10.86	0.03	0.00
50	274.64	174.39	0.96	0.01	0.00
75	321.20	103.61	0.19	0.00	0.00
100	340.15	59.82	0.03	0.00	0.00
125	338.75	36.18	0.04	0.00	0.00

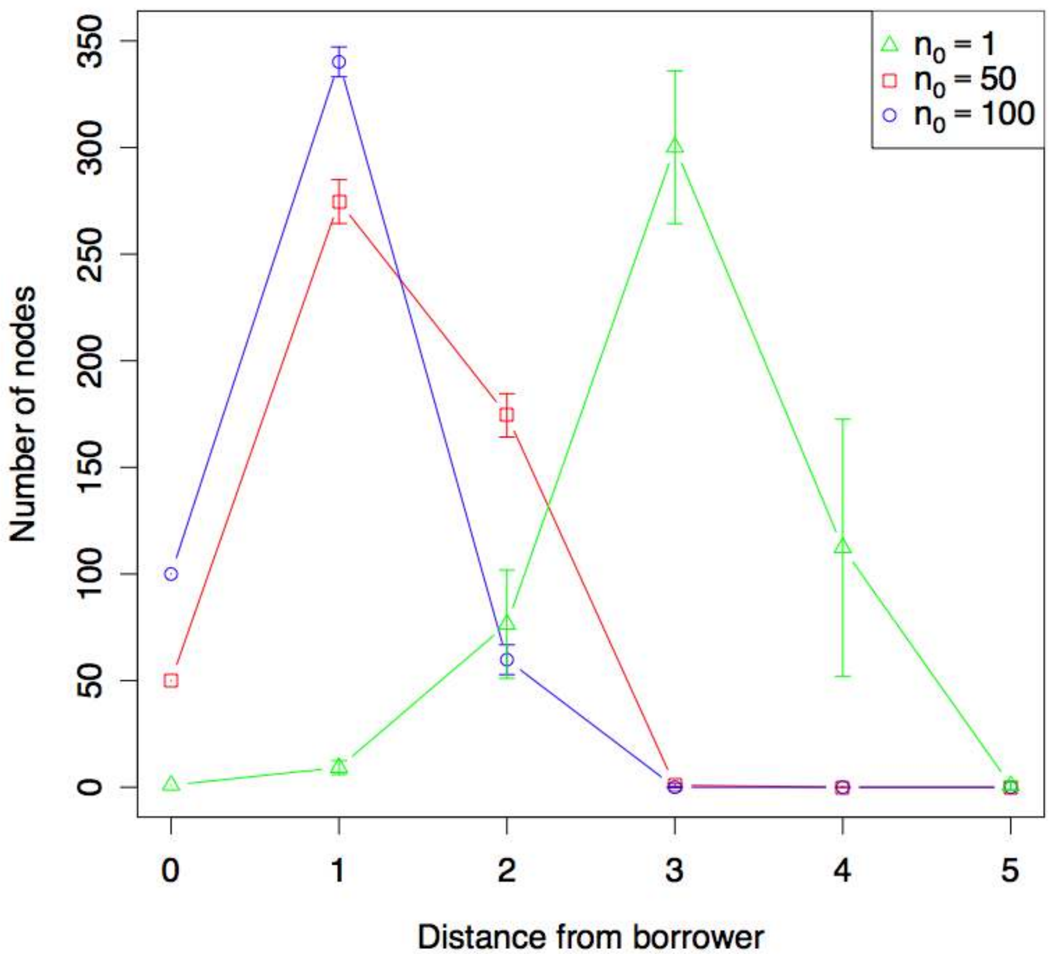


Figure 5.9: Mean values of $n_1 \dots n_5$ in $G(N, p)$ random graphs, $N = 500$ (Table 5.2). Error bars indicate one standard deviation.

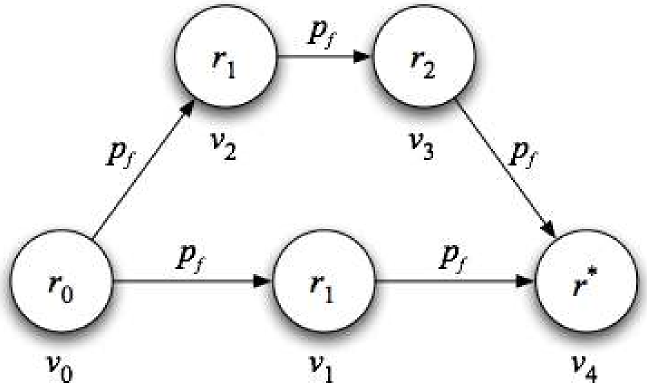


Figure 5.10: Multiple paths between a node and the initial borrower in a graph

Using equation 5.6 with the mean values of $n_1 \dots n_d$ from Table 5.2 will tend to overestimate the true value of $\langle q_f \rangle$ in a graph containing more than one path between a given pair of nodes. Figure 5.10 gives an example of such a graph. Suppose that nodes $v_0 \dots v_3$ have already learned the loanword, and node v_4 is being exposed to tokens of the loanword from both v_1 and v_3. v_4 perceives F tokens from v_1 with probability $r_1 p_f = r_0 p_f{}^2$, while v_4 perceives F tokens from v_3 with probability $r_2 p_f = r_0 p_f{}^3$. The value of r^* will then be some function of these two probabilities, and will depend on the number of tokens that v_1 and v_3 produce for v_4, whether v_1 acquired the loanword before or after v_3, and other such factors. If this graph is converted to a spanning tree which preserves the distances from v_0 for each node (so that v_3 and v_4 are still two links away from v_0), then the edge between v_3 and v_4 will be removed. Thus, according to the tree model, the expected value of r^* is predicted to be exactly $r_0 p_f{}^2$ (since v_4 is linked only to v_1 in the spanning tree); yet the actual value of $\langle r^* \rangle$ in the original graph should be somewhat lower than this, since $p_f{}^3 < p_f{}^2$ and thus v_4 will tend to be exposed to slightly more N tokens from v_3 than from v_1. In other words, the tree model considers only the shortest path from the initial borrower to each node, and ignores the contributions of any other paths, which will tend towards more nativization the longer they are than the shortest path.

To find out the degree to which equation 5.6 overestimates the true value of $\langle q_f \rangle$ in arbitrary graphs, I performed an additional set of simulations of loanword spreading in $G(N, p)$ random graphs, with $N = 500$, $p = 1.5\frac{\log N}{N}$, and $n_0 = 50$ (Figure 5.11), as well as in small-world graphs[12] with similar parameter settings and with $p_{rewire} = 0.1$ (Figure 5.12). The estimated value of $\langle q_f \rangle$ from the model (as indicated by the dashed lines in both graphs) turns out to be slightly larger than the actual values from the simulations, with the greatest difference for intermediate p_f values between 0.5 and 0.8. This shows that the existence of multiple paths from the original borrower causes nativization to be slightly more likely than in a random tree. These multiple paths also have the effect of reducing the variance of q_f, especially for higher values of p_f, as compared to the line model in Figure 5.6. The tree model of transmission thus turns out to be a fairly good approximation for the effects of transmission in random and small-world graphs as well.

5.3 Modeling Japanese adaptation patterns

In the previous section, I have shown that under a simple probabilistic model of loanword borrowing, the transmission of loanwords between the nodes in a random graph causes a nonlinear amplification of the rate of nativization at the individual speaker level. The question now is how to relate this model to the historical patterns of nativization from Chapter 2. Recall the graph from Chapter 2 (reproduced here as Fig-

[12]Small-world graphs were introduced by Watts & Strogatz (1998) as a social network model which can account for both the low average path length and high clustering coefficient seen in real-world social networks (Table 5.1). They are generated by first connecting all of the nodes in a regular lattice structure, with each node having exactly k neighbors (in the simulations I performed here, $k = 2$). Then each node in the graph is visited in turn; for each edge emanating from a particular node, with probability p_{rewire}, it will be disconnected from its destination node and rewired to a randomly-chosen node in the graph. With a certain range of values for p_{rewire}, the networks generated will have low average path lengths, but high clustering coefficients (Watts 1999).

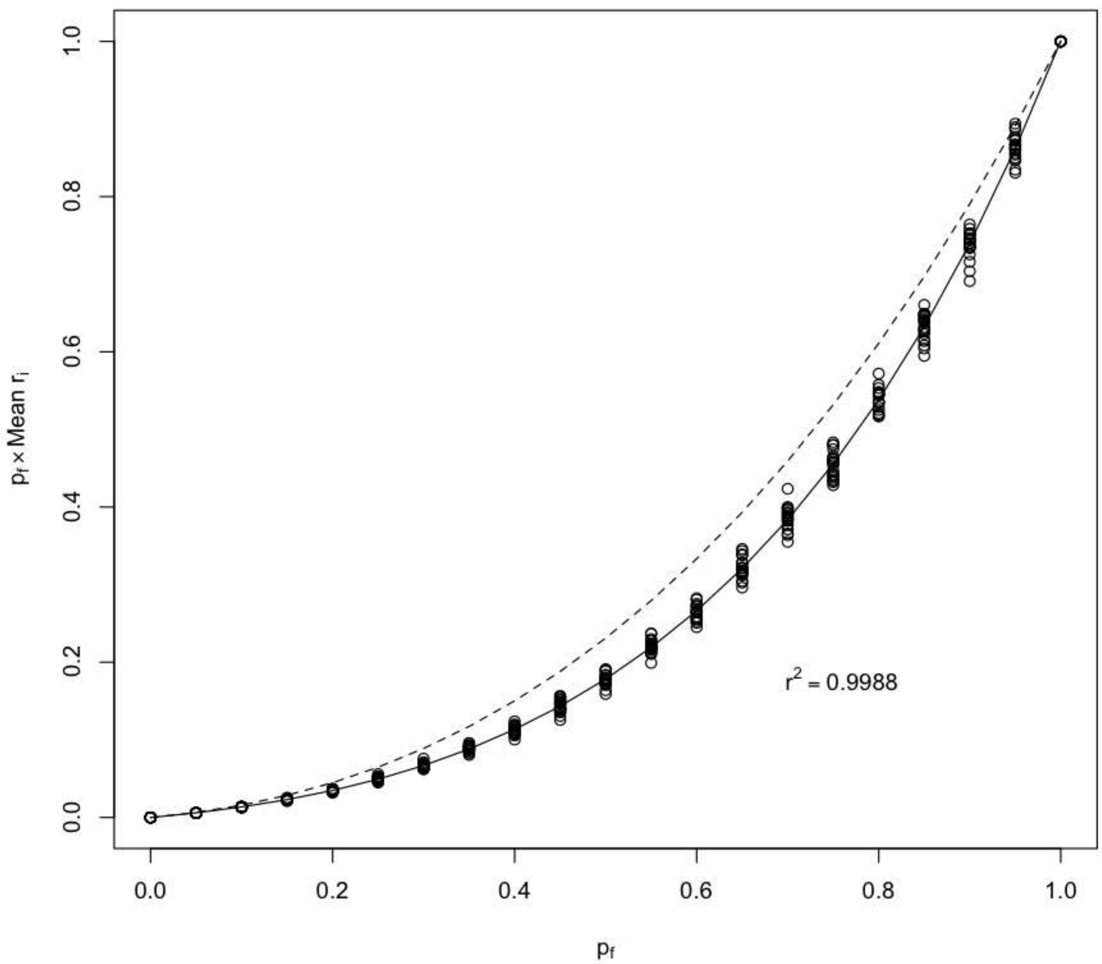


Figure 5.11: Results of simulations of loanword transmission in $G(N,p)$ random graphs, with $N = 500$, $p = 1.5\frac{\log N}{N}$, $n_0 = 50$, and $r_0 = 1$. Twenty networks were generated for each value of p_f tested. Dashed line indicates predicted values of q_f from equation 5.6 with $N = 500$, $n_0 = 50$, and the values of $n_1 \ldots n_5$ from Table 5.2, while the solid line indicates the results of polynomial regression on the simulation results using a fifth-degree polynomial in p_f ($r^2 = 0.9988$).

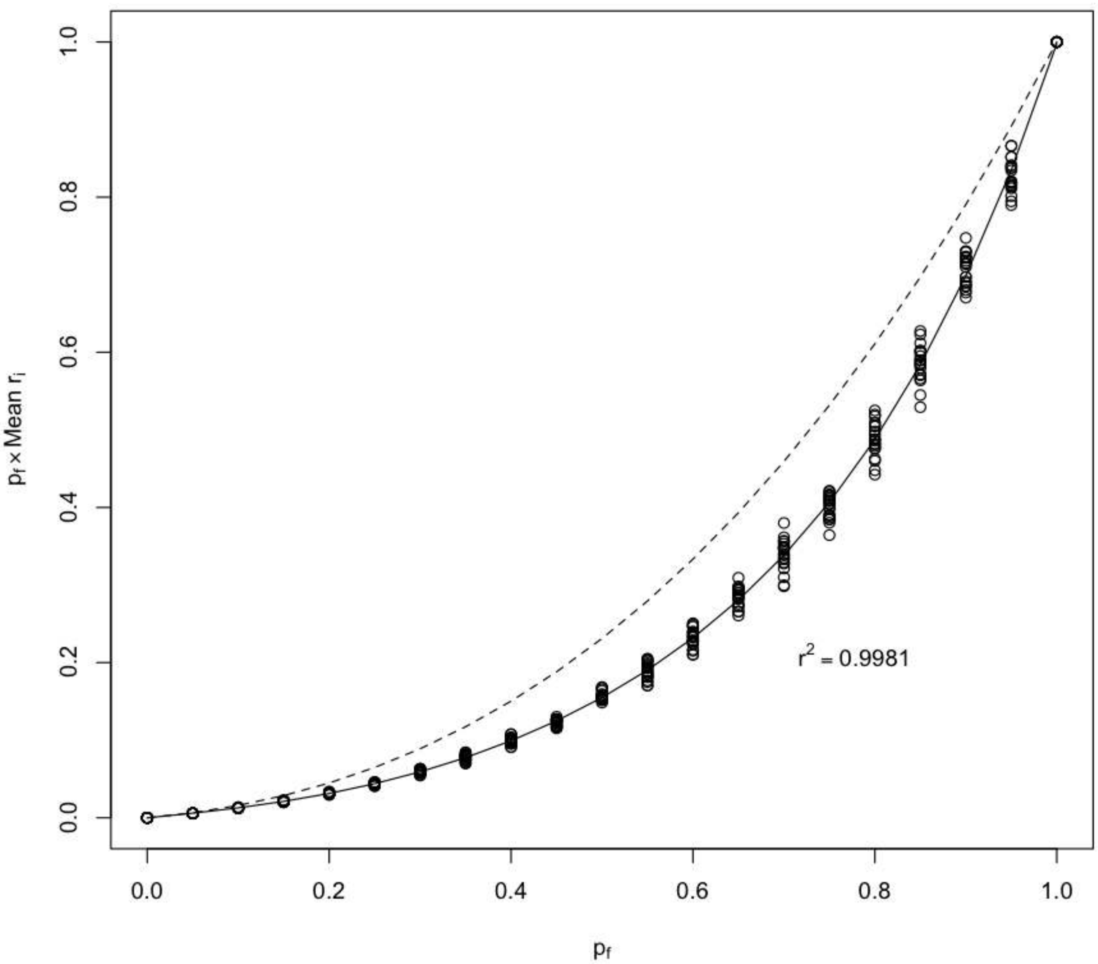


Figure 5.12: Results of simulations of loanword transmission in small-world graphs, with $N = 500$, $n_0 = 50$, $r_0 = 1$, $k = 2$, and $p_{rewire} = 0.1$. The dashed line indicates predicted values of q_f from equation 5.6 using the values of $n_1 \ldots n_5$ from Table 5.2, while the solid line indicates the results of polynomial regression ($r^2 = 0.9981$).

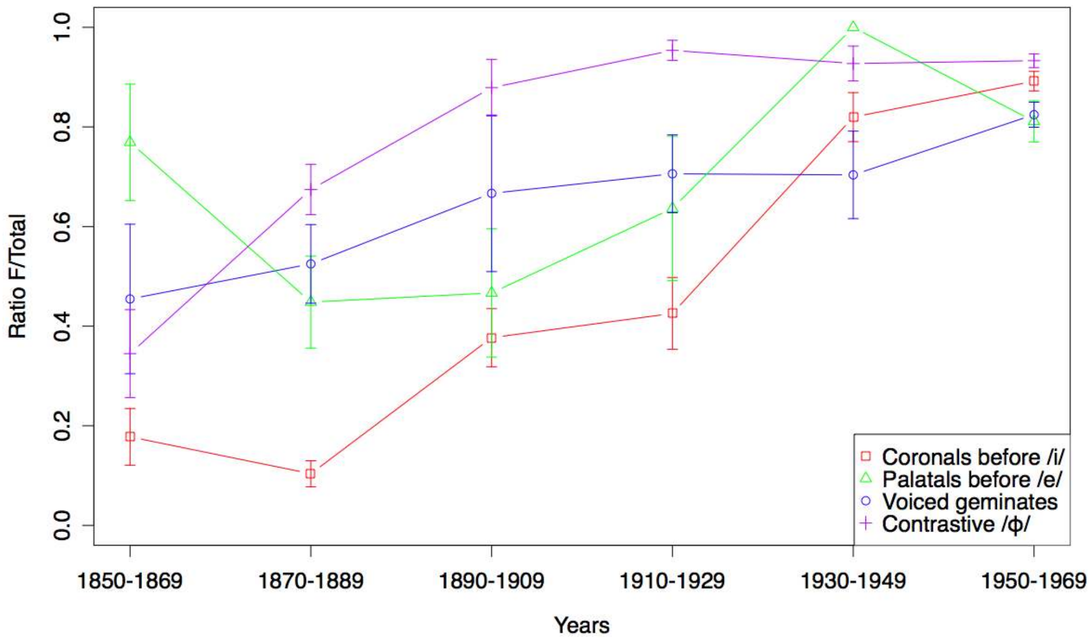


Figure 5.13: Ratios of F tokens to total attested tokens per 20-year period for coronals before /i/, palatals before /e/, voiced geminates, and contrastive /ɸ/

ure 5.13) showing the ratio of F tokens to all tokens attested within a twenty-year span for loanwords first attested in various time periods (1850-1869, 1870-1889, and so on up to 1950-1969). While the four groups of loanwords plotted here (coronals before /i/, palatals before /e/, voiced geminates, and contrastive /ɸ/) all show a general trend towards less nativization for more recently-borrowed words, the data is somewhat noisy, especially for palatals before /e/. Coronals before /i/ show the greatest increase in acceptability over time; they are much more likely to be nativized in loans first attested between 1850-1889 than they are in later borrowings, and they are also more likely to be nativized than loans in the other three categories that are first attested in the same time period. As I discussed in Chapter 2, this suggests that there was some kind of change in the acceptability of TI that took place between 1870 and 1890, after which Japanese speakers were less likely to nativize coronal-/i/ sequences in loanwords.

The Y-axis values in Figure 5.13 can be interpreted as an estimate for the value of $\langle q_f \rangle$ from equation 5.6 for each of these groups of loanwords in each time period. The transmission model can then be used to quantify to what degree these attested rates of nativization in loanwords are due to the action of the initial borrowers, as represented by the parameter r_0 (the probability that the initial borrowers select an F target in producing the loanword), and to what degree are they due to the cumulative effects of misproduction and misperception during transmission, as represented by the parameter p_f (the probability that an F token target is produced and perceived as such during a single act of transmission).[13] First consider the implicit assumption made in previous loanword studies, that nativization is solely due to the initial borrower, and transmission has no effect on the overall likelihood of nativization in the speech community. There are two ways to represent this assumption in the transmission model: either set $p_f = 0$, or $p_f = 1$. Setting $p_f = 0$ would mean that no matter what target form (F or N) any speaker in the network intends to produce, it will always be perceived as the N form by the listener; that is to say, the listener will always nativize whichever variant of the loanword she is exposed to. However, with $p_f = 0$, then equation 5.6 reduces to $\langle q_f \rangle = 0$; in other words, only the N forms of loanwords should ever be attested, and there should never be any variation between F and N forms. Clearly this is not the case historically. Setting $p_f = 1$, on the other hand, would mean that any F targets will always be produced, and perceived, as F tokens. In this case, equation 5.6 reduces to

$$\langle q_f \rangle = \frac{r_0}{N} \sum_k n_k = r_0. \tag{5.7}$$

In other words, the attested ratio of F tokens to total tokens of a loanword should directly reflect the probability that the initial borrowers selected an F target in producing the

[13] I'm only considering here the effects of the two parameters in the model which describe the linguistic behavior of the speakers in the network. The parameters N (number of nodes in the network) and n_k (number of nodes of distance k from an initial borrower) instead relate to the structure of the network, and not directly to the likelihood of nativization on the part of any particular speaker. In the discussion that follows I will assume that these parameters are fixed with the values used in the simulations in the last section ($N = 500$, $n_0 = 50$).

loanword. This would mean that the changes in q_f over time result from corresponding changes in r_0; the Y-axis values in Figure 5.13 would then be an estimate for the value of r_0 for the initial borrowers in each time period. Under this version of the model, the decrease in attested nativization over time would then be caused by an increase in the r_o values for the initial borrowers, or in other words, an increase over time in the likelihood that the initial borrowers will choose an F target for a newly-borrowed word. This increasing tendency to use F tokens as opposed to N tokens among borrowers could plausibly be explained as the result of extragrammatical factors, such as greater exposure to English education among Japanese speakers. However, this result would depend on the unlikely assumption that $p_f = 1$, meaning that historical Japanese speakers were able to perfectly perceive and produce any non-native patterns in a loanword. But if this were the case, then it is not clear why the initial borrowers would ever need to nativize in the first place. If anything, these borrowers, who are in contact with L2, should have more facility with producing and perceiving L2 patterns that don't exist in L1 than do other L1 speakers, most of whom would have been monolingual in Japanese given the historical contact situation.

The transmission model thus shows that with the effect of transmission removed, the historically attested changes in nativization patterns must correlate directly with changes in the borrowers' representations of loanwords, but only if we make the additional, implausible, assumption that transmitting speakers never nativized F tokens themselves. Now let us consider the effects that transmission may have had in determining the values of q_f in each time period. Equation 5.6 can be rewritten as

$$\frac{\langle q_f \rangle}{r_0} = \frac{p_f}{N} \sum_k n_k {p_f}^k = \sum_k \frac{n_k}{N} {p_f}^{k+1}, \tag{5.8}$$

which is a polynomial in p_f that can then be solved for p_f for each attested value of q_f from Figure 5.13 (given appropriate values of r_0, N and n_k). Alternatively, since we know that the q_f curve is a polynomial in p_f, polynomial regression can be performed

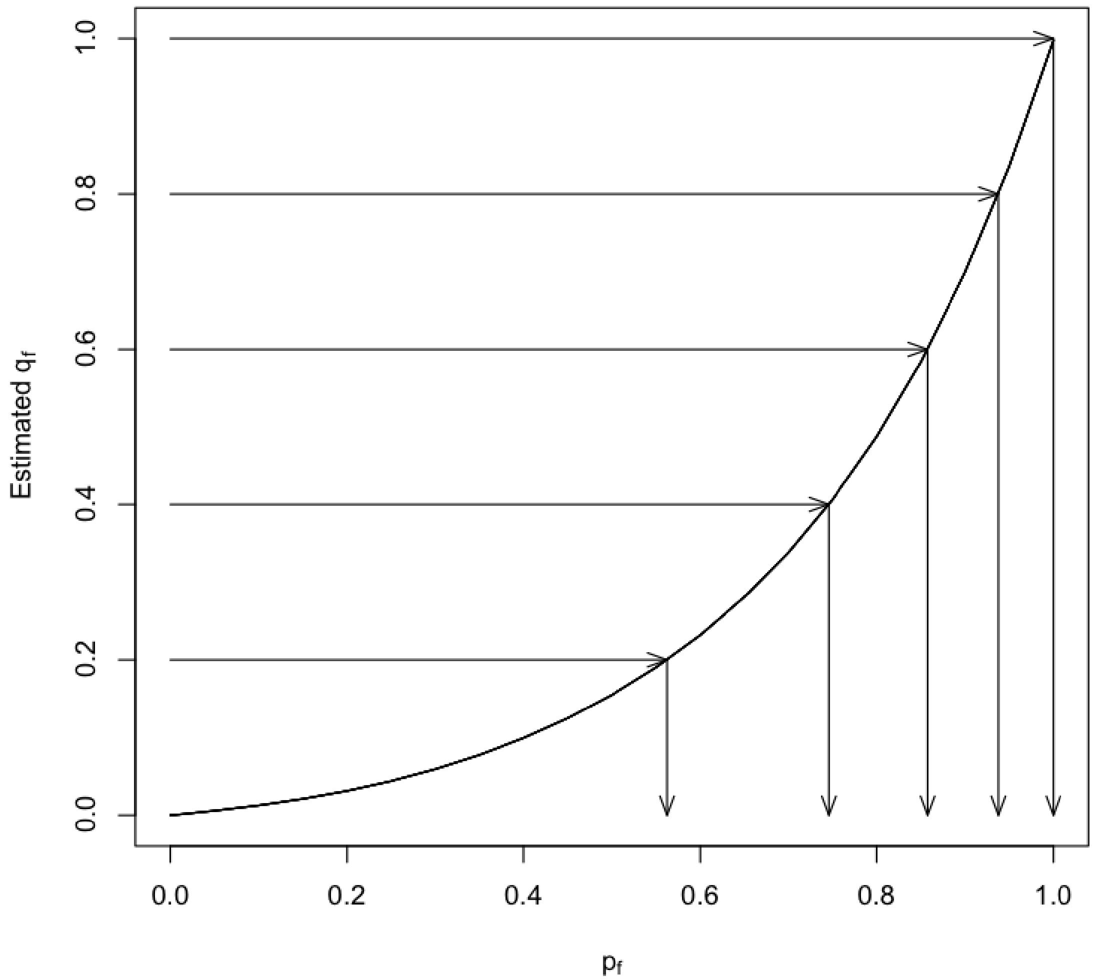


Figure 5.14: Finding p_f values using polynomial regression on the results from Figure 5.12

on the results of the simulations from the previous section to find p_f for a given value of q_f, as illustrated in Figure 5.14. Here the arrows correspond to various values of q_f (0.2, 0.4, 0.6, 0.8, 1.0); the arrows are extended to the right to where they meet the regression curve, then down to the x-axis to find the corresponding value of p_f. The non-linear nature of the effects of transmission is evident from the resulting values of p_f. An increase in the value of q_f from 0 to 0.2 represents a larger change in p_f than an increase in q_f from 0.8 to 1.

Figure 5.15 shows the regressed values of p_f using the small-world network simula-

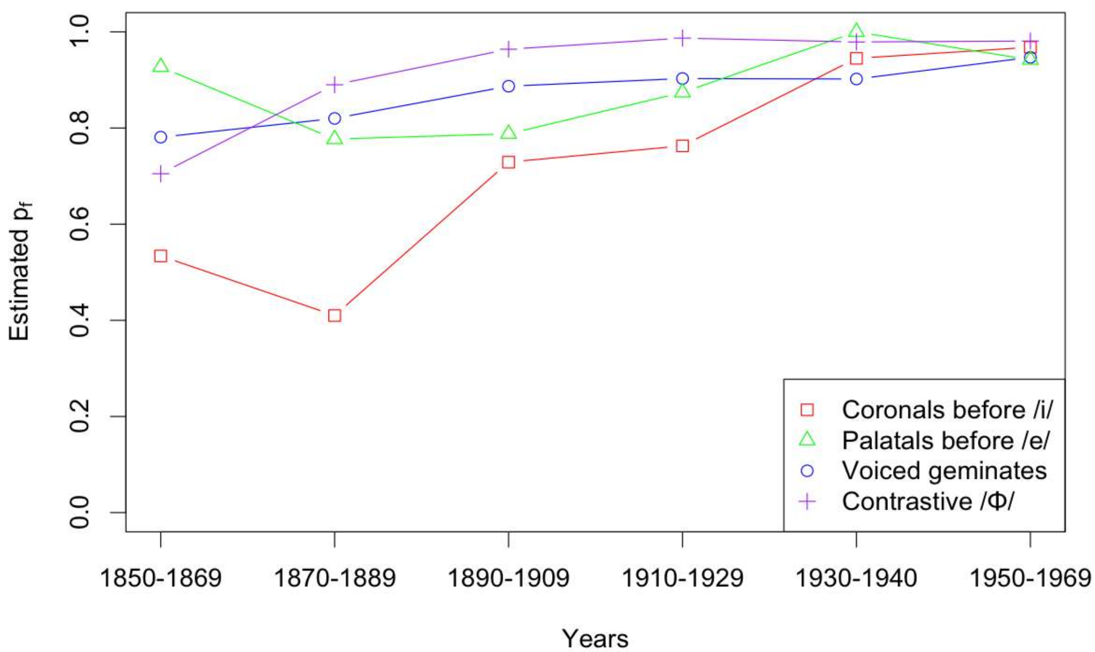


Figure 5.15: Estimated values of p_f for the historical nativization data in Figure 5.13 in a small-world graph, based on the polynomial regression from Figure 5.12, and with $r_0 = 1$.

tions from the previous section, with $r_0 = 1$. Comparing this graph with Figure 5.13, it can be seen that much of the variation in the attested nativization rates can be attributed to the effects of transmission, since the values of p_f for each time period vary over a smaller range than the values of $\langle q_f \rangle$ from Figure 5.13. In particular, in Figure 5.15 the curves for palatals before /e/, voiced geminates, and contrastive /ɸ/ group together, while the curve for coronals before /i/ is different from these first three. This is a typical result; similarly-shaped graphs result using the regressed values from other simulations over a wide range of network parameters. It can be seen from the graph that the predicted p_f values for palatals before /e/, voiced geminates, and contrastive /ɸ/ range from about 0.7-0.8 in the mid-19$^{\text{th}}$ century to about 0.9 in the mid-20$^{\text{th}}$ century.[14] The trend for less nativization over time is still evident in this graph; however, the transmis-

[14] Recall from the end of Chapter 2 that the data point for palatals before /e/ for 1850-1869 is an outlier caused by an unusually high number of words with /ʧe/ first attested during that period.

sion model shows that the relatively high rates of nativization at the beginning stages of contact with English do not necessarily reflect a corresponding inability of Japanese speakers from that time to perceive and produce certain types of non-native phonological patterns. This is because, even if the speakers in the network model have a p_f value of about 0.75, meaning that they are able to correctly perceive and produce F tokens at a rate better than chance, given the average path length of the network in the simulation shown here, the cumulative effects of misperception and misproduction during transmission will still result in the N form of a loanword being attested more frequently than the F form.[15]

The p_f values for loanwords with coronals before /i/, however, show a larger increase over time, even after factoring out the effects of transmission, from approximately 0.5 in 1850-1869 to 0.9 in 1950-1969. This suggests that TI sequences were considerably more difficult for mid-19th century Japanese speakers than the other three non-native patterns, and was actively avoided in the production of TI loanwords. Or, using the adaptation model introduced in Chapter 3, TI was nativized by many mid-19th century Japanese speakers at both the SP (segmental parse) and CS (constraint satisfaction) levels, while nativization of the other three patterns was perhaps an SP-level process only. This is a plausible conclusion, given that coronals before /i/ are more marked in Japanese phonology than the other three non-native sequences being considered here, as was discussed in section 3.2 of Chapter 3. The palatalization of /t/ before /i/ is a highly productive process affecting Japanese verb inflection, while the other three patterns only violate either static generalizations about possible URs, or constraints like *DD which are not violated very often in the native underlying morphology.

[15]This suggests a simple way to empirically test the predictions of the transmission model without having to perform the nearly impossible task of tracking a group of loanwords as they are spread from speaker to speaker in real time. This would be to perform perception and production experiments to find the average value of p_f for various non-native patterns among modern-day Japanese speakers, and then compare the expected value of q_f from equation 5.6 based on the measured value of p_f to the actual rate of nativization among recently attested loanwords.

Note that the TI p_f values begin to rise after about 1890, which correlates with the establishment of mandatory English-language classes in the Japanese educational system. As well, beginning in about 1870, a large number of loans containing TI entered the language, and this would have the effect of decreasing the likelihood of TI being misperceived or misproduced in later loans, as I argued in Chapter 2. These two factors may have been what caused the value of p_f for TI loans to increase to approximately the same level as for loans with ČE and voiced geminates. Of course, the same could be said for the other three groups of loans, that their values of p_f should have also increased after 1870-1890. However, since they were not as marked for mid-19th century Japanese speakers as TI was, they had a relatively high value of p_f to begin with, and thus the increase over time was not as marked as it was with TI.

Thus, the transmission model suggests that, even though mid-19th century loanwords in Japanese show a wide range of nativization rates, much of this range can be attributed to the cumulative effects of misperception and misproduction as these loanwords spread through a network of speakers, and the nativization of these patterns may have been an SP-level process only. However, the attested nativization rate of TI sequences was too low to be solely due to transmission, suggesting that Japanese speakers at this time also had an active phonological constraint at the CS level against TI sequences, in addition to any nativization of TI occurring at the SP level. This constraint became inactive in loanwords after about 1890, due to the increased number of TI words in the Japanese lexicon and English-language education, causing the nativization of TI to become an SP-level process only, just as in the other three non-native patterns.

5.4 Conclusion

In this chapter I have presented a formalized framework for understanding the effects of both on-line adaptation, and transmission of loanwords through a social network, on the resulting form of the established loanword. The model of loanword transmission developed in this chapter predicts that the usual effect of transmission in a social network is to cause a nonlinear amplification of the rate of nativization at the level of individual speakers. I then applied the predictions of this model in explaining the historical adaptation data from Chapter 2 of loanwords with coronals before /i/, palatals before /e/, voiced geminates, and contrastive /ɸ/, showing that much of the variation in the nativization rates of three of these four patterns can be attributed to the effects of transmission. The inferred values of p_f for the four patterns suggest that palatals before /e/, voiced geminates, and contrastive /ɸ/ were correctly produced and perceived by Japanese speakers in all time periods at a rate of about 70-90%. However, even when the effects of transmission are factored out, the data suggest that coronals before /i/ in loanwords were still more likely to be nativized by individual speakers than the other three non-native patterns in the early stages of contact with English, between about 1850-1890. This correlates with the relative markedness of TI compared to the other three patterns in Japanese morphophonology. As well, the increase in p_f for TI loans first attested after 1870 corresponds to a large influx of such loans after this date, caused by the opening of Japan to the Western world. This large number of new loanwords then had the effect of decreasing the rate of misproductions and misperceptions of TI in later loans borrowed after about 1890, to the point where such loans were nativized at approximately the same rate as loans with ČE or voiced geminates.

CHAPTER 6

CONCLUSION

In the previous four chapters I have presented a theoretical framework for modeling loanword borrowing which unifies grammatical and social factors influencing adaptation patterns, as well as providing a coherent account of the sources of variation in these adaptation patterns and how such adaptations can change over generations. I have argued that borrowing is best conceived as a process involving the actions of multiple speakers in a speech community. Borrowing consists of two different subprocesses: adaptation and transmission. Adaptation is the process by which individual L1 speakers map from the L2 source representation of the borrowed word to an L1 representation, while transmission is the cumulative process of L1 speakers learning the adapted loanword from other L1 speakers. These two processes make different demands on the speakers involved, since adaptation involves mapping from a single source form, while transmission needs to resolve conflicts between multiple variants that a speaker may be exposed to. Nativization can take place during either of these processes, resulting in complex patterns of variation which cannot easily be explained with a model that considers the effects of adaptation only.

Adaptation is what much of the previous literature on loanword borrowing has focused on. An L1 speaker performs an adaptation of a new loanword by mapping an L2 surface representation of the word (either phonetic or orthographic) to an L1 phonemic representation. In doing this, the speaker tries to represent the L2 source as faithfully as possible, using the resources made available in the L1 phonology. This tradeoff between faithfulness to the L2 form, and L1 markedness, can be naturally represented in constraint-based theories of phonology like Optimality Theory. In my discussion of three of the adaptation patterns governing Japanese loanwords in Chapter 2, I show that the possible adaptations that are attested for each non-native pattern can be obtained by reranking loanword-specific SB-Correspondence constraints against a fixed ranking of native markedness constraints. An OT grammar can thus be used to delineate a language-internal typology of possible loanword adaptations. However, it is not possible to also account for the cognate effect on velar palatalization in KÆ loanwords (in which KÆ words with transparent cognates in French or German have plain /k/, while KÆ words with no such cognates have palatalized /k^j/) using OT or any other model of a single speaker's phonological competence, without making unrealistic assumptions about each borrowers' knowledge of other languages besides English. Instead, it is necessary to also consider the role of speakers who were involved in transmitting the loanword, specifically, which variant did they prefer after being exposed to both the palatalized and unpalatalized variants from different speakers. The cognate effect thus arises in the competition between the plain and palatalized variants as each spreads through the speech community.

OT-based models also cannot account for frequency and similarity effects on the changes in adaptation patterns over time. This is most clearly seen with the TI loans, in which the non-native sequences /ti/ and /di/ are first attested in high-frequency environments, such as word-finally, and then gradually spread to lower-frequency environments, similar to the lexical diffusion of sound change (Wang 1969, Labov 1981). In Chapters 3 and 4 I consider the implications of these borrowing patterns for the structure of the lexicon and the interaction between the lexicon and the phonological grammar. I argue that, in variants of OT like the Core-Periphery model (Ito & Mester 1999) or the Cophonology model (Inkelas & Zoll 2007) which have been proposed to account for phonological exceptions in loanwords, the lexicon does not have enough of a fine-grained structure to be able to also account for frequency and similarity effects. However, these effects can be explained in connectionist models, like the Triangle Model of Seidenberg & McClelland (1989), which impose a similarity relation over the lexicon, and in which similarity and frequency effects come about as a result of the processing and learning mechanisms of a connectionist network. In Chapter 4 I use an attractor network which is trained on a subset of the Japanese lexicon to examine the adaptation of voiced geminates in loanwords, showing that the relative unacceptability of /bb/ in loans (as compared to /dd/ and /gg/) is due to the unusual status of /p/ in Japanese phonology and the rarity of geminate /p/ in the native lexicon, while the high acceptability of /dd/ is due to the high frequency of /tt/ in the lexicon (including the past tense forms of verbs), as well as the existence of early loans like /beddo/ 'bed'.

Another novel feature of the framework developed in this dissertation is that it also considers the process of loanword transmission. This is in contrast to previous models of borrowing in the generative literature, in which the grammar of a single idealized speaker stands in for the adaptation conventions established over time by the entire speech community, making it difficult to account for synchronic variation in attested loanwords, or diachronic change in adaptation patterns. In Chapter 5 I develop an agent-based model for characterizing the expected effect of transmission on nativization. Assuming a simple probabilistic model of individual speaker adaptation, I find that transmission causes a nonlinear amplification of the rate of nativization in adaptations performed by a single speaker. This nonlinearity increases as the average path length in the social network increases. This means that, even with a relatively low probability of nativization by individual speakers, a randomly chosen speaker in the network can still end up preferring the nativized variant of a loanword over the non-native variant. I then use this model to infer the individual rates of nativization for the adaptation patterns from Chapter 2, given the historically attested rates of nativization. I find that, even after factoring out the expected effect of transmission, the palatalization of coronals in TI loans seems to have taken place at a qualitatively higher rate before 1890 than nativization of other loanword patterns. I suggest that this is evidence that coronal palatalization was a categorical process for many Japanese speakers before 1890, whereas afterwards, perhaps due to increased English-language education, palatalization became more of a low-level phonetic process, being performed by speakers at rates comparable to the nativization of other loanword patterns.

There are a variety of developments of this basic theoretical framework for loanword borrowing which I plan to pursue in future research, as I will now discuss in the remainder of this chapter.

6.1 Phonological grammars and connectionist models of lexical processing

I will be extending the connectionist model of loanword adaptation from Chapter 4 to account for a variety of adaptation patterns seen in loanwords in Japanese and other languages. The current model uses phonological representations only, so I will work on adding orthographic representations as well, in a manner similar to Plaut et al. (1996) and Harm & Seidenberg (1999), so that I can also account for the spelling adaptations seen in Japanese loanwords, and how these interact with phonetically-based adaptations. The challenge here is that there are three different writing systems (*kanji*, *hiragana*, and *katakana*) used in Japanese; as well, Romanization is taught in the context of English education, and many adaptation patterns (such as the adaptation of English /ə/) reflect the spelling of the source word in English. It will be quite difficult to design an orthographic representation layer for a connectionist network that can efficiently represent all four of these scripts. Given the large number of *kanji* characters in Japanese, a localist representation, in which there is one unit for each possible character, would probably be too large in terms of the number of units for the resulting network to be trainable in a reasonable amount of time, and would have poor generalization capabilities from orthography to phonology, since the radicals present in the *kanji* characters would not be directly represented. Thus a distributed representation for Japanese orthography would be more practical. One possibility would be to use a feature-based representation where the features represent the presence of a single stroke in a specific position and orientation (for example, a vertical stroke on the left of the character). Another would be to use random vectors for each character, where the vectors are chosen so that their similarity to the vectors for other characters corresponds to the characters' visual similarity.

Another possible extension to the connectionist model of loanword adaptation is to add in the semantic level of representation from the TM framework. This would make it possible to account for morphological alternations, such as the present and past tense of verbs. In the current model, the past tense alternation was represented by training the network on the phonological forms only of both verb forms. Since there were no semantic representations used, the network treated these items as separate lexical entries, and thus could not represent the fact that the two verb forms are morphologically related. By training a network on both phonological and semantic representations, with the present and past tense forms differing in a systematic way (for example, using a [past] feature), then the network would learn a constraint governing the value of the [past] feature and the corresponding presence of gemination in the phonological form. In other words, the network would be able to learn that the gemination that occurs in many past tense forms is a morphophonological alternation, and not simply a static distribution over phonological forms in the lexicon. This would make it possible to examine the effect that these alternations have on adaptation patterns. I suspect that such a network, like historical Japanese speakers, would be more likely to nativize phonotactic sequences such as /ti/ which are avoided in morphophonological alternations, as compared to sequences such as /ʧe/ which also do not occur in the lexicon but are not generated underlyingly in the morphology.

I also plan on collecting more historical data, from such sources as newspaper archives and Meiji-era novels, showing how various adaptation patterns are attested in the late 19th and early 20th centuries. I will focus on the palatalization of coronal obstruents before /i/, as this adaptation pattern is the most interesting in terms of showing the effects of phonological similarity and frequency, as well as the palatalization of velars before source /æ/, which shows the effects of competition during transmission. This data will also allow me to estimate the token frequencies of loanwords in this time period, making it possible to more accurately model the lexical knowledge of Meiji-era Japanese speakers.

On a more theoretical level, I wish to explore further the connections between Optimality Theory and Harmonic Grammar, and the Triangle Model for lexical processing proposed in Seidenberg & McClelland (1989). The advantage of TM and other connectionist models is their ability to account for similarity and frequency effects in irregular or variant forms using the same processing mechanisms as for regular forms. I believe that it should be possible to analyze the weights in a trained TM network to derive an HG grammar which would be a higher-level description of the network's behavior. This derivation would make explicit some of the constraints which the network is implementing during processing, thus making it easier to understand how the network operates. Going in the other direction, I also want to explore how the representations and constraint interactions proposed in recent phonological theories can be used to constrain the possible network architectures in the TM framework. This would allow for building connectionist models that are not only psychologically plausible, but linguistically plausible as well. I believe that such combined models will be able to account for both the patterns of regularity and productivity that are seen in phonological systems, as well as the variation in the use of such systems that is seen at the level of the individual speaker.

6.2 Agent-based simulations of language contact and language change

I plan on developing further the agent-based model of loanword transmission from Chapter 5 in several ways. I will test some of the empirical predictions of the current model by doing perception and production experiments on native Japanese speakers using nonce words with non-native phonotactic sequences such as [si], and comparing the rate at which these sequences are nativized during perception and production with the attested nativization rate of these same sequences in recent loanwords in a text corpus. The prediction based on the transmission model is that, if speakers nativize with a probability $p_n = 1 - p_f$, then the rate of the non-native variant occurring in text should be a polynomial function of p_f, with the terms of the polynomial reflecting the expected distribution of path lengths in the network from initial borrowers to other speakers.

I will also explore the effects of extending the model by allowing for more linguistically realistic behavior at the level of the individual speaker, for example by having the agents in the network have different tendencies to nativize or not, or by having the agents evaluate loanword forms they are exposed to using a stochastic grammar like Harmonic Grammar, or a TM network like the attractor network from Chapter 4, rather than a simple random choice between forms. This will also make it possible to look at the interaction between the frequency and similarity effects on adaptations and the likelihood of non-native variants being transmitted. Yet another possible change that could be made to the model at the level of the individual speaker is to assign a social status value to each agent, and have agents prefer variants that are produced by higher-status agents in the network, in a manner similar to the simulations performed by Nettle (1999a,b). This extension to the model would then be compared to a corresponding neutral model of loanword variation, in which agents simply choose loanword variants on the basis of which variants are used by their neighbors in the network, without regards to the neighbors' social status value, to determine the influence that higher status agents have on determining adaptation patterns (cf. Shalizi 2007). These extensions will make deriving an analytical model significantly more difficult, and thus much harder to characterize the behavior of the models for various values of their parameters. Perhaps techniques from evolutionary game theory (Young 1998, Gintis 2000) will be useful for this task, as exemplified by Jäger's (2007) application of game theory to the typology of word order.

Finally, the most challenging task will be to extend the model to account for changes in adaptation patterns taking place over several generations. The current transmission model looks only at the spread of a single loanword, which can occur over a relatively short period of time. In order to account for changes in adaptation patterns over longer periods of time, it will be necessary to extend the model of the social network itself by allowing for new agents to be added to the graph over time and old agents to be removed, simulating the birth and death of members of the speech community. Perhaps in the course of language acquisition, the new agents in the network, based on their exposure to loanwords being transmitted in the network, as well as their degree of English-language education, end up setting the value of p_f at a higher value than their predecessors. This would result in the qualitative pattern discussed at the end of Chapters 2 and 5, where the likelihood of nativization decreases over time. More generally, to account for the effect of bilingualism on nativization (Haugen 1950), it will be necessary to model increasing bilingualism in the social network, for example by gradually increasing the number of contacts that agents have with L2 speakers outside the network, or by making it more likely that newly-added agents will be attached in the network near L2 speakers or L1/L2 bilinguals.

Using this extended model would allow for the exploration of the interactions between individual speakers' phonological competences and their expectations for the adaptations that other speakers in the network will produce, providing some clues as to how to relate synchronic variation and social structure to diachronic change, as in the *actuation problem* posed by Weinreich et al. (1968). This model would also have clear connections to Evolutionary Phonology (Blevins 2004) and similar approaches (Ohala 1981, 1993, Lindblom 1986) in which sound changes arise from the cumulative perception and production biases of individual speakers. The difference would be that the social network model proposed here would be able to account for synchronic variation, as well as diachronic change, using the same transmission mechanism taking place between different types of agents over different timescales. Synchronic variation would result from transmission between neighbors in the network, while diachronic change would result from transmission between parent and child agents.

BIBLIOGRAPHY

Akamatsu, Tsutomu. 1997. *Japanese phonetics: Theory and practice*. Lincom Europa.

Albert, Réka & Albert-László Barabási. 1999. Topology of evolving networks: Local events and universality. *Physical Review Letters* 85(24). 5234–5237.

Amaral, Luís A. Nunes, A. Scala, Marc Barthelemy & H. Eugene Stanley. 2000. Classes of small-world networks. *Proceedings of the National Academy of Sciences (PNAS)* 97(21). 11149–11152.

Anttila, Arto. 2002. Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory* 20(1). 1–42.

Anttila, Arto & Young-mee Yu Cho. 1998. Variation and change in Optimality Theory. *Lingua* 104(1–2). 31–56.

Arakawa, Sōbē (ed.). 1977. *Gairaigo jiten [Loanword dictionary]*. Tokyo: Kadokawa Shoten.

Barabási, Albert-László & Réka Albert. 1999. Emergence of scaling in random networks. *Science* 286(5439). 509–512.

Beech, Hannah. 1998. The day the phones died. *Time* 151(9).

Bergs, Alexander. 2005. *Social networks and historical sociolinguistics: Studies in morphosyntactic variation in the Paston letters*. Mouton de Gruyter.

Bishop, Christopher M. 1995. *Neural networks for pattern recognition*. Oxford University Press.

Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.

Bloch, Bernard. 1950. Studies in colloquial Japanese IV: Phonemics. *Language* 26(1). 86–125.

Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart, and Winston.

Boersma, Paul & Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistc Inquiry* 32. 45–86.

Butters, Ronald R. 1995. Important clarification. *American Speech* 70(1). 20.

Bybee, Joan L. & Dan L. Slobin. 1982. Rules and schemas in the development and use of the English past tense. *Language* 58. 265–289.

Carrington, Peter J., John Scott & Stanley Wasserman (eds.). 2005. *Models and methods in social network analysis*. Cambridge University Press.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.

Christiansen, Morten H., Joseph Allen & Mark S. Seidenberg. 1998. Learning to segment speech using multiple cues: A connectionist model. *Language and Cognitive Processes* 13. 221–268.

Christiansen, Morten H. & Nick Chater (eds.). 2001. *Connectionist psycholinguistics: The state of the art*. Westport, CT: Ablex Publishing.

Christiansen, Morten H., Rick Dale & Florencia Reali. In press. Connectionist explorations of multiple-cue integration in syntax acquisition. In Scott P. Johnson (ed.), *Neoconstructivism: The new science of cognitive development*. Oxford University Press.

Clements, G. N. & Elizabeth V. Hume. 1995. The internal organization of speech sounds. In Goldsmith (1995), 245–306.

Coltheart, Max. 1981. Disorders of reading and their implications for models of normal reading. *Visible Language* 15. 245–286.

Coltheart, Max, Brent Curtis, Paul Atkins & Michael Haller. 1993. Models of reading aloud: Dual-route and Parallel-Distributed-Processing approaches. *Psychological Review* 100. 589–608.

Cox, Trevor F. & Michael A. A. Cox. 2001. *Multidimensional scaling*. Chapman and Hall/CRC, 2nd edn.

Crawford, Clifford. 2004. The Core-Periphery model of the lexicon and phonotactic constraints on loanwords in Japanese. In *Proceedings of the 2004 Linguistic Society of Korea International Conference*, 97–107.

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

Crawford, Clifford. 2008. An evolutionary account of loanword-induced sound change in Japanese. *Penn Working Papers in Linguistics* 14(1). 57–70.

Dell, Gary S. 2000. Commentary: Lexical representation, counting, and connectionism. In Michael B. Broe & Janet B. Pierrehumbert (eds.), *Papers in laboratory phonology V: Acquisition and the lexicon*.

Dell, Gary S. & Jean K. Gordon. 2003. Neighbors in the lexicon: Friends or foes? In Niels O. Schiller & Antje S. Meyer (eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities*, vol. 6, Phonology and Phonetics, 9–37. Mouton de Gruyter.

Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier & Jacques Mehler. 1999. Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25(6). 1568–1578.

Eckert, Penelope. 2000. *Linguistic variation as social practice: The linguistic construction of identity in Belten High*. Blackwell Publishers Ltd.

Electronic Dictionary Research and Development Group. 2003. `JMDICT` [Japanese-English dictionary]. Tech. rep., Monash University, School of Computer Science and Software Engineering. URL `http://www.csse.monash.edu.au/~jwb/j_jmdict.html`.

Elman, Jeffrey L. 1991. Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning* 7. 195–225.

Elman, Jeffrey L., Elizabeth A. Bates, Mark H. Johnson, Annette Karmiloff-Smith, Domenico Parisi & Kim Plunkett. 1996. *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Epstein, Joshua M. 2007. *Generative social science: Studies in agent-based computational modeling*. Princeton, NJ: Princeton University Press.

Erdős, Paul & Alfréd Rényi. 1960. The evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* 5. 17–61.

Frellesvig, Bjarke. 1995. *A case study in diachronic phonology: The Japanese onbin sound changes*. Aarhus University Press.

Fries, Charles C. & Kenneth L. Pike. 1949. Coexistent phonemic systems. *Language* 25. 29–50.

Ghiselin, Michael T. 1974. A radical solution to the species problem. *Systematic Zoology* 23(4). 536–544.

Ghiselin, Michael T. 1997. *Metaphysics and the origin of species*. State University of New York Press.

Gibbons, Alan. 1985. *Algorithmic graph theory*. Cambridge University Press.

Gintis, Herbert. 2000. *Game theory evolving: A problem-centered introduction to modeling strategic behavior*. Princeton University Press.

Goldrick, Matthew. 2007. Constraint interaction: A lingua franca for stochastic theories of language. *MIT Working Papers in Linguistics* 53. 95–114.

Goldsmith, John A. (ed.). 1995. *The handbook of phonological theory*. Oxford: Blackwell Publishers Ltd.

Granovetter, Mark S. 1973. The strength of weak ties. *The American Journal of Sociology* 78(6). 1360–1380.

Gusfield, Dan. 1997. *Algorithms on strings, trees, and sequences: Computer science and computational biology*. Cambridge University Press. Reprinted with corrections, 1999.

Hale, Mark. 2007. *Historical linguistics: Theory and method*. Blackwell Publishing.

Hale, Mark & Charles Reiss. 2000. "Substance abuse" and "Dysfunctionalism": Current trends in phonology. *Linguistic Inquiry* 31(1). 157–169.

Hale, Mark & Charles Reiss. 2001. Phonology as cognition. In Noel Burton-Roberts, Philip Carr & Gerard Docherty (eds.), *Phonological knowledge: Conceptual and empirical issues*, 161–184. Oxford University Press.

Hamano, Shoko. 1998. *The sound-symbolic system of Japanese*. Stanford, CA: CSLI Publications.

Hamming, Richard W. 1950. Error detecting and error correcting codes. *Bell System Technical Journal* 26(2). 147–160.

Hancin-Bhatt, Barbara & Anita Govindjee. 1999. A computational model of feature competition in L2 phonology. In Peter Broeder & Jaap Murre (eds.), *Language and thought in development: Cross-linguistic studies*, 145–161. Tübingen: Gunter Narr Verlag.

Hanson, Stephen José & David J. Burr. 1990. What connectionist models learn: Learning and representation in connectionist networks. *Behavioral and Brain Sciences* 13. 471–518.

Harary, Frank. 1972. *Graph theory*. Addison-Wesley, 3rd edn.

Hare, Mary & Jeffrey L. Elman. 1995. Learning and morphological change. *Cognition* 56(1). 61–98.

Harm, Michael W. & Mark S. Seidenberg. 1999. Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review* 106(3). 491–528.

Haugen, Einar. 1950. The analysis of linguistic borrowing. *Language* 26(2). 210–231.

Hayes, Bruce & Donca Steriade. 2004. The phonetic bases of phonological markedness. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically based phonology*, 1–33. Cambridge University Press.

Hebb, Donald O. 1949. *The organization of behavior*. New York: Wiley.

Higa, Masanori. 1970. The sociolinguistic significance of borrowed words in the Japanese spoken in Hawaii. *University of Hawaii Working Papers in Linguistics* 2. 125–140.

Hornik, Kurt, Maxwell Stinchcombe & Halbert White. 1989. Multilayer feedforward networks are universal approximators. *Neural Networks* 2(5). 359–366.

Hyman, Larry. 1970. The role of borrowing in the justification of phonological grammars. *Studies in African Linguistics* 1(1). 1–48.

Ichikawa, Sanki. 1930. The pronunciation of English loan-words in Japanese. In N. Bøgholm, Aage Brusendorff & C. A. Bodelsen (eds.), *A grammatical miscellany offered to Otto Jespersen on his seventieth birthday*, 179–190. Copenhagen: Levin and Munksgaard.

Inkelas, Sharon & Cheryl Zoll. 2007. Is grammar dependence real? A comparison between cophonological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics* 45(1). 133–171.

Ito, Junko. 1990. Prosodic minimality in Japanese. *Chicago Linguistic Society (CLS)* 26. 213–239.

Ito, Junko & R. Armin Mester. 1986. Phonological voicing in Japanese. *Linguistic Inquiry* 17. 49–73.

Ito, Junko & R. Armin Mester. 1995. Japanese phonology. In Goldsmith (1995).

Ito, Junko & R. Armin Mester. 1999. The phonological lexicon. In Natsuko Tsujimura (ed.), *A handbook of Japanese linguistics*. Oxford: Blackwell Publishers Ltd.

Ito, Junko & R. Armin Mester. 2003. *Japanese morphophonemics: Markedness and word structure*. Cambridge, MA: MIT Press.

Iverson, Gregory K. & Ahrong Lee. 2006. Perception of contrast in Korean loanword adaptation. *Korean Linguistics* 13. 49–87.

Iverson, Paul, Patricia K. Kuhl, Reiko Akahane-Yamada, Eugen Diesch, Yoh'ich Tohkura, Andreas Ketterman & Claudia Siebert. 2003. A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition* 87. B47–B58.

Jacobs, Haike & Carlos Gussenhoven. 2000. Loan phonology: Perception, salience, the lexicon, and OT. In Joost Dekkers, Frank Reinoud Hugo van der Leeuw & Jeroen Maarten van de Weijer (eds.), *Optimality Theory: Phonology, syntax, and acquisition*, 193–210. Oxford University Press.

Jäger, Gerhard. 2007. Evolutionary game theory and typology: A case study. *Language* 83(1). 74–109.

James, Ioan. 1999. *Topologies and uniformities*. Springer.

Kager, Rene. 1999. *Optimailty theory*. Cambridge University Press.

Kamermans, Michiel. 2008. Frequency lists. Available from the Monash Nihonogo FTP Archive: `http://ftp.monash.edu.au/pub/nihongo/00INDEX.html`.

Kang, Yoonjung. 2003. Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology* 20. 219–273.

Katayama, Motoko. 1998. *Optimality theory and Japanese loanword phonology*. Ph.D. thesis, University of California, Santa Cruz.

Kawahara, Shigeto. 2006. A faithfulness ranking projected from a perceptibility scale: The case of [+voice] in Japanese. *Language* 82(3). 536–574.

Kawahara, Shigeto, Kohei Nishimura & Hajime Ono. 2003. Unveiling the unmarkedness of Sino-Japanese. In *Japanese/Korean linguistics*, vol. 11, 140–151. Stanford, CA: CSLI.

Keating, Patricia & Aditi Lahiri. 1993. Fronted velars, palatalized velars, and palatals. *Phonetica* 50. 73–101.

Kenstowicz, Michael & Atiwong Suchato. 2006. Issues in loanword adaptation: A case study from Thai. *Lingua* 116(7). 921–949.

Kossinets, Gueorgi & Duncan Watts. 2006. Empirical analysis of an evolving social network. *Science* 311. 88–90.

Kruskal, Joseph B. & Myron Wish. 1978. *Multidimensional scaling*, Quantitative Applications in the Social Sciences. Beverly Hills, CA: Sage Publications, Inc.

Kuroda, S.-Y. 2002. Rendaku. In Noriko M. Akatsuka & Susan Straus (eds.), *Japanese/Korean linguistics*, vol. 10, 337–350. CSLI Publications.

Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: University of Philadelphia Press.

Labov, William. 1981. Resolving the Neogrammarian controversy. *Language* 57. 267–308.

Labrune, Laurence. 2002. The prosodic structure of simple abbreviated loanwords in Japanese: A constraint-based account. URL `http://roa.rutgers.edu/`. ROA–532.

LaCharité, Darlene & Carole Paradis. 2005. Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic Inquiry* 36. 223–258.

Lahiri, Aditi & Vincent Evers. 1991. Palatalization and coronality. In Carole Paradis & Jean-François Prunet (eds.), *The special status of coronals: Internal and external evidence*, vol. 2, Phonetics and Phonology, 79–100. San Diego: Academic Press, Inc.

Lange, Roland A. 1973. *The phonology of eighth-century Japanese*. Tokyo: Sophia University.

Lass, Roger. 1980. *On explaining language change*. Cambridge University Press.

Leci, Mark & Meoni Poon. 2004. Resolving consonant clusters: An OT analysis of English loanwords into Cantonese. *University of Washington Working Papers in Linguistics* 23. 103–118.

Legendre, Géraldine, Yoshiro Miyata & Paul Smolensky. 1990. Harmonic Grammar—A formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations. In *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*, 884–891. Cambridge, MA: Lawrence Erlbaum.

Levenshtein, Vladimir I. 1966. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady* 10. 707–710.

Liljeros, Fredrik, Christofer R. Edling, Luís A. Nunes Amaral, H. Eugene Stanley & Yvonne Åberg. 2001. The web of human sexual contacts. *Nature* 401(6840). 907–908.

Lindblom, Björn. 1986. Phonetic universals in vowel systems. In J. J. Ohala & J. J. Jaeger (eds.), *Experimental phonology*. Academic Press.

Loveday, Leo J. 1996. *Language contact in Japan: A socio-linguistic history*. Oxford University Press.

Lovins, Julie. 1975. *Loanwords and the phonological structure of Japanese*. Indiana University Linguistics Club.

Luce, Paul A. & David B. Pisoni. 1998. Recognizing spoken words: The neighborhood activation model. *Ear and Hearing* XX. 19–36.

Martin, Samuel E. 1952. *Morphophonemics of standard colloquial Japanese*. Ph.D. thesis, Yale University.

Martin, Samuel E. 1987. *The Japanese language through time*. Yale University Press.

Mayr, Ernst. 1982. *The growth of biological thought: Diversity, evolution, and inheritance*. Cambridge, MA: Belknap Press.

McCarthy, John J. 1988. Feature geometry and dependency: A review. *Phonetica* 45. 84–108.

McCawley, James D. 1968. *The phonological component of a grammar of Japanese*. The Hague: Mouton.

McClelland, James L. 1994. The interaction of nature and nurture in development: A parallel distributed processing perspective. In Paul Hertelson, Paul Eelen & Géry d'Ydewalle (eds.), *International perspectives on psychological science*, vol. I: Leading Themes, 57–88. Lawrence Erlbaum Associates. Keynote Lectures presented at the XXV International Congress of Psychology, 1992.

McClelland, James L. & Jeffrey L. Elman. 1986. The TRACE model of speech perception. *Cognitive Psychology* 18. 1–86.

Milgram, Stanley. 1967. The Small World problem. *Psychology Today* 2. 60–67.

Milroy, Lesley. 1987. *Language and social networks*. Blackwell, 2nd edn.

Milroy, Lesley. 2002. Social networks. In J. K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.), *The handbook of language variation and change*, 549–572. Blackwell.

Minsky, Marvin & Seymour Papert. 1969. *Perceptrons*. Cambridge, MA: MIT Press.

Miyake, Marc Hideo. 2003. *Old Japanese: A phonetic reconstruction*. RoutledgeCurzon.

Murray, Stephen O. 1993. Network determination of linguistic variables? *American Speech* 68(2). 161–177.

Navarro, Gonzalo. 2001. A guided tour to approximate string matching. *ACM Computing Surveys* 33(1). 31–88.

Nettle, Daniel. 1999a. *Linguistic diversity*. Oxford University Press.

Nettle, Daniel. 1999b. Using Social Impact Theory to simulate language change. *Lingua* 108. 95–117.

Newman, John. 1997. *Coursebook in feature geometry*. Lincom Europa.

Newman, Mark E. J. 2003. Random graphs as models of networks. In Stefan Bornholdt & Heinz Georg Schuster (eds.), *Handbook of graphs and networks: From the genome to the internet*, 35–68. WILEY-VCH.

Nishimura, Kohei. 2003. Lyman's Law in loanwords. Tokyo: University of Tokyo, ms.

Nowak, Martin. 2006. *Evolutionary dynamics: Exploring the equations of life*. Cambridge, MA: Belknap Press.

Oh, Mira. 1996. Linguistic input to loanword phonology. *Studies in Phonetics, Phonology and Morphology* 2. 117–126.

Ohala, John. 1981. The listener as a source of sound change. *Chicago Linguistic Society (CLS)* 17. 178–203.

Ohala, John. 1983. The origin of sound patterns in vocal tract constraints. In Peter F. MacNeilage (ed.), *The production of speech*, 189–216. New York: Springer-Verlag.

Ohala, John. 1993. Sound change as nature's speech perception experiment. *Speech Communication* 13. 155–161.

Ohso, Mieko. 1971. *A phonological study of some English loan words in Japanese*. Master's thesis, The Ohio State University.

Ota, Mitsuhiko. 2004. The learnability of the stratified phonological lexicon. *Journal of Japanese Linguistics* 20. URL `http://roa.rutgers.edu/`. ROA–668.

Paradis, Carole & Darlene LaCharité. 1997. Preservation and minimality in loanword adaptation. *Journal of Linguistics* 33. 379–430.

Pater, Joe. 2008. Optimization and linguistic typology. URL `http://roa.rutgers.edu/`. Ms., University of Massachusetts, Amherst. ROA–982.

Paul, Hermann. 1880/1889. *Prinzipien der sprachgeschichte [Principles of the history of language]*. New York: MacMillan and Co. Translated from the second edition of the original by H. A. Strong.

Pearsall, Judy & Bill Trumble (eds.). 2002. *The Oxford English reference dictionary*. Oxford, England: Oxford University Press, 2nd edn.

Peperkamp, Sharon. 2004. A psycholinguistic theory of loanword adaptations. *Berkeley Linguistics Society (BLS)* 30.

Peperkamp, Sharon & Emmanuel Dupoux. 2003. Loanword adaptations: Three problems for phonology (and a psycholinguistic solution). *International Congress of Phonetic Sciences* 15. 367–370.

Pierrehumbert, Janet B. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In Joan Bybee & Paul Hopper (eds.), *Frequency effects and emergent grammar*. Amsterdam: John Benjamins.

Pinker, Steven. 1991. Rules of language. *Science* 253. 530–535.

Pinker, Steven & Alan Prince. 1988. On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition* 28(1–2). 73–193.

Plate, Tony A. 1994. *Distributed representations and nested compositional structure*. Ph.D. thesis, University of Toronto.

Plaut, David C. 2001. A connectionist approach to word reading and acquired dyslexia: Extension to sequential processing. In Christiansen & Chater (2001), chap. 8, 244–278.

Plaut, David C., James L. McClelland, Mark S. Seidenberg & Karalyn Patterson. 1996. Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review* 103(1). 56–115.

Poplack, Shana, David Sankoff & Chris Miller. 1988. The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26(293). 47–104.

Poser, William. 1990. Evidence for foot structure in Japanese. *Language* 66. 78–105.

Prince, Alan & Paul Smolensky. 1993. Optimality theory: Constraint interaction in generative grammar. Tech. rep., Rutgers University and University of Colorado at Boulder. URL http://roa.rutgers.edu/. ROA–537 version, 2002.

Prince, Alan & Paul Smolensky. 2006. Optimality: From neural networks to Universal Grammar. In Smolensky & Legendre (2006), chap. 4, 123–143.

R Development Core Team. 2006. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/. ISBN 3-900051-07-0.

Redner, Sidney. 1998. How popular is your paper? An empirical study of the citation distribution. *The European Physical Journal B* 4(2). 131–134.

Reed, Russell D. & Robert J. Marks. 1999. *Neural smithing: Supervised learning in feedforward artificial networks*. Cambridge, MA: MIT Press.

Rice, Keren. 1997. Japanese NC clusters and the redundancy of postnasal voicing. *Linguistic Inquiry* 28. 541–551.

Ripley, Brian D. 1996. *Pattern recognition and neural networks*. Cambridge University Press.

Rohde, Douglas. 2000. Lens: The light, efficient network simulator. URL http://tedlab.mit.edu:16080/~dr/Lens/.

Rosch, Eleanor. 1973. Natural categories. *Cognitive Psychology* 4(3). 328–350.

Rosch, Eleanor & Carolyn B. Mervis. 1975. Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology* 7(4). 573–605.

Rosenblatt, Frank. 1958. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review* 65(6). 386–408.

Rosenblatt, Frank. 1962. *Principles of neurodynamics: Perceptrons and the theory of brain mechanisms*. Spartan Books.

Rumelhart, David E., Geoffrey E. Hinton & Ronald J. Williams. 1986. Learning internal representations by error propagation. In David E. Rumelhart, James L. McClelland & the PDP Research Group (eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*, vol. 1, chap. 8, 318–362. Cambridge, MA: MIT Press.

Saciuk, Bohdan. 1969. The stratal division of the lexicon. *Papers in Linguistics* 1. 464–532.

Sagey, Elizabeth. 1990. *The representation of features in non-linear phonology: The articulatory node hierarchy*. New York: Garland.

Sarle, Warren S. 2002. Neural network FAQ. URL `ftp://ftp.sas.com/pub/neural/FAQ.html`. Posted monthly to the newsgroup `comp.ai.neural-nets`.

Scott, John. 2000. *Social network analysis: A handbook*. Newberry Park, CA: Sage, 2nd edn.

Searcóid, Mícheál Ó. 2006. *Metric spaces*. Springer-Verlag.

Seidenberg, Mark S. 2005. Connectionist models of word reading. *Current Directions in Psychological Science* 14(5). 238–242.

Seidenberg, Mark S. & James L. McClelland. 1989. A distributed, developmental model of word recognition and naming. *Psychological Review* 96(4). 523–568.

Shalizi, Cosma Rohilla. 2007. Social media as windows on the social life of the mind. URL `http://arxiv.org/abs/0710.4911`. Forthcoming in the proceedings of the 2008 AAAI symposium on “Social Information Processing”.

Shepard, Roger N. 1987. Toward a universal law of generalization for psychological science. *Science* 237(4820). 1317–1323.

Shibata, Takeshi. 1976. Sekai no naka no nihongo [Japanese in the world]. In *Nihongo to kokugogaku [Japanese and Japanese linguistics]* (Iwanami-kōza Nihongo 1), 1–30. Tokyo: Iwanami.

Shibatani, Masayoshi. 1999. *The languages of Japan*. Cambridge University Press.

Shinohara, Shigeto. 1996. The roles of the syllable and the mora in Japanese: The adaptation of French words. *Cahiers de Linguistique Asie Orientale* 25(1). 87–112.

Silverman, Daniel. 1992. Multiple scansions in loanword phonology: Evidence from Cantonese. *Phonology* 9. 289–328.

Silverstein, Michael. 1972. Chinook Jargon: Language contact and the problem of multi-level generative systems, I. *Language* 48(2). 378–406.

Simon, Herbert A. 1955. On a class of skew distribution functions. *Biometrika* 42. 425–440.

Smith, Jennifer L. 2006. Loan phonology is not all perception: Evidence from Japanese loan doublets. In Timothy J. Vance & Kimberly A. Jones (eds.), *Japanese/Korean linguistics*, vol. 14, 63–74. CSLI.

Smolensky, Paul. 1990. Tensor product variable binding and the representation of symbolic structures in connectionist networks. *Artificial Intelligence* 46. 159–216.

Smolensky, Paul. 2006a. Formalizing the principles I: Representation and processing in the mind/brain. In Smolensky & Legendre (2006), chap. 5, 147–206.

Smolensky, Paul. 2006b. Optimization in neural networks: Harmony maximization. In Smolensky & Legendre (2006), chap. 9, 345–392.

Smolensky, Paul. 2006c. Tensor product representations: Formal foundations. In Smolensky & Legendre (2006), chap. 8, 271–344.

Smolensky, Paul & Géraldine Legendre (eds.). 2006. *The harmonic mind: From neural computation to Optimality-Theoretic grammar*. Cambridge, MA: MIT Press.

Stampe, David L. 1973. *A dissertation on natural phonology*. Ph.D. thesis, University of Chicago. Distributed 1979 by Indiana University Linguistics Club, Bloomington, IN.

Steedman, Mark. 2001. Connectionist sentence processing in perspective. In Christiansen & Chater (2001), chap. 11, 348–372.

Steen, Lynn Arthur & J. Arthur Seebach, Jr. 1995. *Counterexamples in topology*. Dover.

Stemberger, Joseph Paul. 2004. Neighbourhood effects on error rates in speech production. *Brain and Language* 90. 413–422.

Stockall, Linnaea & Alec Marantz. 2006. A single route, full decomposition model of morphological complexity: MEG evidence. *The Mental Lexicon* 1(1).

Tateishi, Koichi. 1990. Phonology of Sino-Japanese morphemes. *University of Massachusetts Occasional Papers in Linguistics* 13. 209–235.

Thomason, Sarah Grey & Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. University of California Press.

Travers, Jeffrey & Stanley Milgram. 1969. An experimental study of the Small World problem. *Sociometry* 32(4). 425–443.

Tsuchida, Ayako. 1995. English loans in Japanese. *Working Papers of the Cornell Phonetics Laboratory* 10. 145–164.

Valente, Thomas. 1995. *Network models of the diffusion of innovation*. Cresskill, NJ: Hampton Press.

Vance, Timothy J. 1985. Glottal stops in Japanese. *Journal of Asian Culture* 9. 21–35.

Vance, Timothy J. 1987. *An introduction to Japanese phonology*. Albany, NY: State University of New York Press.

Venables, William N. & Brian D. Ripley. 2002. *Modern applied statistics with S*. New York: Springer, 4th edn. URL `http://www.stats.ox.ac.uk/pub/MASS4`. ISBN 0-387-95457-0.

Vendelin, Inga & Sharon Peperkamp. 2006. The influence of orthography on loanword adaptations. *Lingua* 116(7). 996–1007.

Vitevitch, Michael S. & Paul A. Luce. 1999. Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language* 40. 374–408.

Wang, William S.-Y. 1969. Competing changes as a cause of residue. *Language* 45. 9–25.

Wang, William S.-Y. (ed.). 1977. *The lexicon in phonological change*. The Hague: Mouton.

Wasserman, Stanley & Katherine Faust. 1994. *Social networks analysis: Methods and applications*. Cambridge University Press.

Watts, Duncan. 1999. *Small worlds: The dynamics of networks between order and randomness*. Princeton, NJ: Princeton University Press.

Watts, Duncan & Steven Strogatz. 1998. Collective dynamics of ‘small-world’ networks. *Nature* 393. 440–442.

Weinreich, Uriel. 1953. *Languages in contact: Findings and problems*. New York: Linguistic Circle of New York.

Weinreich, Uriel. 1957. On the description of phonic interference. *Word* 13(1). 1–11.

Weinreich, Uriel, William Labov & Marvin Herzog. 1968. Empirical foundations for a theory of language change. In Winfred Lehmann & Yakov Malkiel (eds.), *Directions for historical linguistics*, 97–195. Austin, TX: University of Texas Press.

Williams, Ronald J. & Jing Peng. 1990. An efficient gradient-based algorithm for on-line training of recurrent network trajectories. *Neural Compuation* 2. 490–501.

Yip, Moira. 1993. Cantonese loanword phonology and Optimality Theory. *Journal of East Asian Linguistics* 2(3). 261–293.

Yip, Moira. 2006. The symbiosis between perception and grammar in loanword phonology. *Lingua* 116(7). 950–975.

Young, H. Peyton. 1998. *Individual strategy and social structure: An evolutionary theory of institutions*. Princeton, NJ: Princeton University Press.

Yule, G. Udny. 1925. A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S. *Philosophical Transactions of the Royal Society of London, Series B* 213. 21–87.

Zipf, George K. 1935. *The psychobiology of language*. Houghton-Mifflin.