Title

The Historical Development of Initial Accent in Trimoraic Nouns in Kyoto Japanese

Author

Angeles, Andrew

Publication Date

2019

UNIVERSITY OF CALIFORNIA

SANTA CRUZ

THE HISTORICAL DEVELOPMENT OF INITIAL ACCENT IN TRIMORAIC NOUNS IN KYOTO JAPANESE

A thesis submitted in partial satisfaction
of the requirements for the degree of

MASTER OF ARTS

in

LINGUISTICS

by

Andrew Angeles

September 2019

The thesis of Andrew Angeles
is approved:

Professor Junko Ito, Chair

Associate Professor Ryan Bennett

Associate Professor Grant McGuire

Quentin Williams
Acting Vice Provost and Dean of Graduate Studies

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

The Historical Development of Initial Accent in Trimoraic Nouns in Kyoto Japanese

Andrew Angeles

This paper discusses the pathway, motivating factors, and mechanism of the historical development of the initial accent pattern in trimoraic nouns in Kyoto Japanese, tracing its development from being one of the least common patterns in Early Middle Japanese (ca. 12th century AD) to becoming the most common pattern for accented nouns in present-day Modern Kyoto Japanese.

Accent, defined as a pitch fall from high pitch to low pitch (Haraguchi 1999), if it is present, may in principle fall on any mora in a Kyoto Japanese word, as illustrated by the example in (1).

(1) Possible accent patterns in Modern Kyoto Japanese

a.	I’noti	‘life’	H’LL	initial accent
b.	OTO’ko	‘man’	HH’L	second accent
c.	kaBU’to	‘helmet’	LH’L	second accent
d.	osaK'i	‘the future’	LLF	final accent

Examination of the frequencies of each accent pattern reveals that possible accent locations are not equally prevalent. Instead, there is a clear preference for initial accent, with 68% of trimoraic accented native nouns having initial accent (Yoshida and Zamma 2001) and 73% of trimoraic accented nouns across all lexical strata having initial accent (calculated from raw data from Sugito 1995). Despite this present-day prevalence of

the initial accent pattern, Kyoto Japanese accented trimoraic nouns did not always exhibit this preference. In the accentual system of Early Middle Japanese, the earliest recorded ancestral accentual system of the Kyoto Japanese branch, the initial accent pattern accounted for only about 6% of accented trimoraic nouns (calculated from Sugito 1995).

Previous investigations (e.g. Hata and Hasegawa 1988, Hasegawa and Hata 1995, Kawakami 1995 as cited by Nakai 2001, Matsumori 1999, 2008, and Shimabukuro 2007) have proposed accounts for what is likely to have happened in the developmental pathway from Early Middle Japanese to Modern Kyoto Japanese, suggesting that several accent classes merged into the initial accent pattern or otherwise merged into patterns which would eventually come to have initial accent. One pathway described in these accounts is as follows. A pitch rise enhancement process proposed by Kawakami (1995) occurred, whereby an L tone immediately preceding an H tone was lowered in order to enhance the contrast between them, causing any L tones preceding the lowered L tone to rise to H. Thus, for LLH, the following changes would occur: LLH > MLH > HLH. By this process, the Early Middle Japanese forms LLL, LLH, and LLF would gain the initial H tone needed to eventually merge into initial accent (HLL), becoming HHL, HLH, and HLF respectively. HLH and HLF became HLL due to a pressure for words to have only one peak ("culminativity") (Nakai 2001, Shimabukuro 2007). HHL then underwent leftward kernel shift, shifting the lowering kernel on the second mora to the first mora, yielding HLL (Nakai 2001).

Although these accounts propose plausible changes to account for the differences between Early Middle Japanese and Modern Kyoto Japanese, these accounts generally do not focus on what factors were the seeds of change which contributed to the occurrence of the proposed changes, with the exception of Hata and Hasegawa (1988) and Hasegawa and Hata (1995), who attribute the rightward shift of accent to peak delay.

In this paper, I argue that the changes proposed in previous work can be attributed to pattern frequencies that cause learners to induce and rerank constraints in response to these frequencies, changing the grammar in learning generations and leading to the mergers of several accentual classes and thus the development of initial accent as the most common accented pattern in trimoraic nouns. Accordingly, in the development of Modern Kyoto Japanese, pattern frequencies are the seeds of change. Using the pitch rise enhancement process described above as a starting point, I propose four stages of development, with all but one (Stage 3) building on pattern frequencies created by the previous stage. The proposal is framed in the framework of Optimality Theory (Prince and Smolensky 1993/2004), which allows for the precise specification of the state of the grammar at every stage in terms of competing pressures whose influence changes as constraints are reranked through the generations.

Pattern frequencies in pre-pitch rise enhanced nouns in the accentual system of Early Middle Japanese already show a weak preference for word-initial H tones. This weak preference in conjunction with the increased frequencies of word-initial H from Kawakami's pitch rise enhancement process causes learners to induce INITIAL-H, a constraint which requires words to begin with an H tone, in Stage 1 and promote it to a higher rank. In Stage 2, INITIAL-H will have consequences for culminativity, preferring HLH > HLL instead of HLH > LLH. The constraint CULMINATIVITY becomes active because of relatively low frequencies of words with two peaks, causing learners to promote the constraint. In Stage 3, peak delay causes the change LHH > LLH. While this change does not directly lead to an increase in initial accent, as the new LLH forms do not become initial-accented, it nonetheless plays a role in reducing the amount of multiply-linked H tones in the accentual system, serving as the catalyst for Stage 4. In Stage 4, the change HHL > HLL occurs due to the promotion of NOMULTILINK-H, a constraint prohibiting H tones from being linked to multiple moras (Ito and Mester 2018), on the basis of the reduced prevalence of multiply-linked H tones resulting from peak delay in the previous stage.

Finally, a test of the initial Early Middle Japanese data using Hayes and Wilson (2008)'s UCLA Phonotactic Learner induces a constraint similar to the INITIAL-H constraint proposed here, suggesting that learners can induce constraints and vary their strengths based on pattern frequencies in the input data.

The contribution of the present investigation, therefore, is twofold. First, the present study advances our understanding of the development of the modern-day prevalence of initial accent in Modern Kyoto Japanese trimoraic accented nouns. Second, the present study advances our understanding of what kinds of factors can influence the reranking and promotion of constraints, both in Japanese and beyond. As one such factor, the

present study proposes that constraints can be reranked on the basis of the surface frequencies of patterns present in the input learners receive.

ACKNOWLEDGMENTS

I would like to express my deepest appreciation to my committee Junko Ito, Ryan Bennett, and Grant McGuire, who provided advice, suggestions, and support during the many stages of this project and without whose guidance this project would not have been possible. I would also like to extend my gratitude to Armin Mester, who offered suggestions which played a central role in the development of the analysis presented here. Thanks are also due to Nick Kalivoda for his thoughtful suggestions and conversations about the project; Haruo Kubozono, who offered his thoughts and suggestions for places to look for Japanese-language resources; Larry Hyman, who gave helpful comments and questions on my presentation of this work at PHREND 2019; Gorka Elordieta, who provided interesting comments and conversation on this project and a related Modern Kyoto Japanese project; Matt Wagers, who led the Fall 2018 Pre-290 Research Seminar preview meetings; the participants of Pre-290 and the Winter 2019 290 Research Seminar – Jérémie Beauchamp, Netta Ben-Meir, Richard Bibbs, Ben Eischens, Andrew Hedding, Lisa Hofmann, and Nick Van Handel; and the participants of the Phlunch meeting where I presented an earlier version of my work. All have provided valuable insight and feedback on various stages of the project and helped me develop and present a better project.

Finally, I want to thank my family, most especially my parents, Juan and Alessandra Angeles, for their support and encouragement every step of the way. I am truly grateful.

1 Introduction

Japanese nouns may be accented or unaccented, where “accent” is defined as a pitch fall from high to low. Accented nouns show a pitch fall, while unaccented nouns lack it. Accented nouns can be further divided into nouns with unpredictable (underlyingly specified) accent and nouns with predictable (rule-governed) accent (Martin 1952, who referred to these classes as “athematic accent” and “thematic accent” respectively). Words with predictable accent, which include verbs and adjectives, have accent which predictably falls on the antepenultimate or penultimate syllable. Words with unpredictable accent have accent whose location must be lexically specified. Nouns in the native lexical stratum are typically considered to have underlyingly specified accent, allowing accent to occur, if at all, on any mora. However, while accent can, in principle, fall on any mora in a noun, the possible accent patterns are not equally prevalent. In Tokyo Japanese and Modern Kyoto Japanese, the majority of trimoraic nouns are unaccented, and if we restrict our view to only accented trimoraic nouns, the majority have initial accent (Kubozono 2008 for Tokyo Japanese; Sugito 1995 and Yoshida and Zamma 2001 for Modern Kyoto Japanese). Despite this finding, this prevalence of initial accent in trimoraic accented nouns does not seem to have been the case historically as recorded in the *Ruijumyogisho*, a Chinese-Japanese dictionary which recorded the accent patterns of words in 12th century Early Middle Japanese, where initial accent seems to have been relatively uncommon. The present investigation, focusing on trimoraic nouns in the native lexicon, attempts to account for the numerical discrepancy between Early

Middle Japanese (EMJ) on the one hand and Modern Kyoto Japanese on the other hand by means of phonologization of a pitch rise enhancement process that lowers an L-toned mora immediately preceding an H-toned mora, causing any L-toned moras which precede the lowered L-toned mora to raise to H. This process causes an increase in the frequency of word-initial H tones in the input data received by learners, which in turn causes learners to induce a constraint INITIAL-H, which requires that words begin with an H tone. The induction of this constraint is the first in a series of developments consisting of 1) induction of INITIAL-H, 2) promotion of a restriction prohibiting multiple H tones in a word (culminativity), 3) peak delay, and 4) promotion of a restriction against H tones being linked to multiple moras in a word, with the results of each stage having consequences for stages further down the line. Altogether, this chain of developments will lead to the merger of several accentual patterns in Early Middle Japanese (EMJ) into the initial accented pattern, accounting for its present-day prevalence, and will ultimately lead to the Modern Kyoto accent system in trimoraic nouns as presently observed. The developmental pathway proposed herein is presented within the framework of Optimality Theory (Prince and Smolensky 1993/2004). The advantage of understanding this development in Optimality Theory lies in its ability to provide precise specifications of the state of the grammar at every developmental stage and allow us to understand the grammatical history of Kyoto Japanese in terms of competing pressures whose influence changes as constraints are reranked through the generations. These grammars are specified in terms of constraint hierarchies involving violable constraints which can be reranked between stages, due to, I argue, the changing landscape of inputs in terms of accentual pattern type frequencies. Understanding the development of the grammar in this way accounts for the fact that grammars at earlier stages in the development have consequences for grammars at later stages, the kinds of changes that occurred and which patterns underwent change, and the reasons unchanged patterns remained unchanged at any stage of the development.

Central to the present account is that learning generations *induce* a constraint based on increased surface frequencies of initial H tones in the inputs that they receive from the parent generation. The induced INITIAL-H constraint is a language-specific constraint. I argue that if the frequencies of a certain pattern – here, word-initial H in Kyoto Japanese – tend to some degree toward favoring that pattern, such language-specific constraints can be induced, if not immediately (by the first learning generation exposed to a weak trend toward favoring the relevant pattern), eventually (after iterated learning across generations has strengthened the trend). Examination of the kinds of constraints induced by Hayes and Wilson (2008)'s UCLA Phonotactic Learner, a maximum entropy-based learner, reveals that even at the base frequencies of word-initial H in Early Middle Japanese words from the *Ruijumyogisho* before pitch rise enhancement occurs, the Phonotactic Learner nonetheless induced a weak constraint with the same function as INITIAL-H, suggesting that the seed of change already existed in the 12th century system recorded in the *Ruijumyogisho*. As the frequency of word-initial H increases due to the pitch rise enhancement process, so too does the strength of the INITIAL-H constraint, until it is strong enough to influence members of the learning generation to begin the developmental pathway toward initial accent in Modern Kyoto by applying a grammar which includes INITIAL-H to the inputs it receives.

This paper is organized as follows. Section 2 presents an overview of the Tokyo and Modern Kyoto pitch accent systems. Section 3 discusses the prevalence of initial accent in bimoraic and trimoraic nouns in Tokyo and Modern Kyoto and the discrepancy in this prevalence between Early Middle Japanese and Tokyo and Modern Kyoto. Reference is made to Tokyo in Sections 2 and 3 to aid in framing the problem. In the remaining sections, I focus on the development of Modern Kyoto. Section 4 discusses the *Ruijumyogisho* and its interpretation and presents an overview of the developmental pathway from Early Middle Japanese to Modern Kyoto. Section 5 presents the main proposal of the present account, detailing each of the four stages of the developmental pathway and accounting for it in Optimality Theory. Section 6 discusses a learning model which can account for the phonologization of the increased prevalence of initial H as well as a maximum entropy-based simulation of tonotactic learning. Section 7 concludes.

2 Overview of Tokyo Japanese and Modern Kyoto Japanese Pitch Accent

Japanese is a pitch accent language which makes use of changes in pitch to mark the presence – words may either have or lack accent – and location of accent in words (see Kawahara 2015 for an overview and references cited therein). Pitch accent languages may also be described as a subset of the class of tone languages, using fewer tone contrasts and sparser underlying specifications (Yip 2002, Hyman 2009). In tonal terms, Tokyo Japanese (henceforth "Tokyo") and Modern Kyoto Japanese (henceforth "Modern Kyoto") "accent" can be characterized as an HL[1] pitch fall (following Pierrehumbert and Beckman 1988). Unlike stress languages like English and Dutch (Beckman 1986) and a subset of pitch accent languages like Sanskrit, Ancient Greek, and North Gyeongsang Korean (Ito and Mester 2016), which require a prominence ("accent") in all lexical words, accent is not required in another subset of pitch accent languages, including many dialects of Japanese, such as Tokyo and Modern Kyoto[2] (Haraguchi 1999), Somali (Hyman 1981), and dialects of Northern Bizkaian Basque (Hualde, Elordieta, Gaminde, and Smiljanić 2002).

There are many Japanese dialects, which are often divided into three major groups, with 16 subdivisions in total shared across the groups. These groups are 1) the Ryukyuan dialects, which have three subdivisions, 2) the Western Japanese dialects, which have six subdivisions, and 3) the Eastern Japanese dialects, which have seven subdivisions (Shibatani 1990). Further divisions can also be made in terms of prefecture (of which there are 47; e.g. the Tokyo dialect vs. the Kyoto dialect, spoken in the Tokyo and Kyoto prefectures respectively), city (e.g. the Kagoshima City

---

[1] The following tone abbreviations are used in this paper: H = high tone, M = mid tone, L = low tone, F = falling tone, [HL] = falling tone when reference to the internal structure of the falling tone is required, R = rising tone. When diacritics are used, the following conventions are used: H = acute (e.g. á), L = grave (e.g. à), and F = circumflex (e.g. â).

[2] The Modern Kyoto Japanese accent system is very similar to the Osaka Japanese accent system, both being dialects of the Kansai region. There are only minor dialectal variations between the two systems (Yoshida and Zamma 2001). According to Yoshida and Zamma, "most of the Kyoto data fits well with that of the Osaka dialect." As such, data drawn from both dialects are used in this paper and collectively referred to as "Modern Kyoto Japanese."

dialect), or geographical distinction (mainland vs. island; e.g. Kagoshima City dialect on the mainland Kyushu Island vs. Koshikijima Island dialects; see Kubozono 2019). The pitch accent systems of these various dialects differ, often substantially, from dialect to dialect. Tokyo and Modern Kyoto are no exception, with dialects differing on parameters such as how many possible accentual patterns the dialect exhibits and how many tonal registers the dialect has (if any) (Haraguchi 1999, Uwano 1999).

Tokyo and Modern Kyoto are examples of "multi-pattern" accent systems in the terms of Uwano (1999). In multi-pattern systems, the number of possible accentual patterns increases as words become longer. Tokyo exhibits as many as n + 1 accentual patterns in words consisting of n moras, with the "+ 1" referring to an unaccented pattern. Hence, a word consisting of three moras will have one of up to four accentual patterns (3 + 1 = 4): it will either be unaccented or have an accentual prominence (the H of an HL complex) on any one of its three moras. These four accentual possibilities are exemplified in (1) below for Tokyo. Syllables in capital letters indicate association with an H tone; syllables in lowercase letters indicate association with an L tone. The nominative marker *ga* is provided to distinguish between the unaccented word in (1a) and the final accented word in (1d), which would otherwise be identical. Tone schematics are also provided abstracting away from segmental content. Accent location is indicated by an apostrophe (') placed immediately after the syllable bearing the accentual H. All data presented in examples (1) through (3) are from Sugito (1995) and Haraguchi (1999).

(1) Possible accentual patterns for 3 mora words in Tokyo

a. kiTUNE-GA[3]	‘fox (Nom.)’	LHH-H	unaccented
b. I’noti-ga	‘life (Nom.)’	H’LL-L	initial accent
c. koKO’ro-ga	‘heart (Nom.)’	LH’L-L	second accent
d. oTOKO’-ga	‘man (Nom.)’	LHH’-L	final accent

Modern Kyoto exhibits up to 2n + 1 accentual patterns in words consisting of n moras (Haraguchi 1999). Thus, a word consisting of three moras in Modern Kyoto will have one of up to seven possible accentual patterns (2*3 + 1 = 7).[4] The extra three patterns, compared to the four possibilities available to three mora words in n + 1 Tokyo, come from the fact that Modern Kyoto distinguishes not only the presence and location of accent, but whether a word has one or another overall accentual pattern. This distinction is called “register” by Uwano (1999).

A register is a tonal pattern which characterizes an accentual unit as a whole (Uwano 1999). Register as defined by Uwano is only a characteristic of multi-pattern

---

[3] All romanizations in this paper use the phonemic Kunrei-shiki romanization system. In Modern Japanese, certain phonemes exhibit allophony when preceding high vowels. The phonemes /t/ and /d/ have the allophones [tʃ] and [dʒ] before the high front vowel /i/ and the allophones [ts] and [dz] before the high back vowel /ɯ/ (romanized <u> in Kunrei-shiki). The phonemes /s/ and /z/ have the allophones [ʃ] and [dʒ] before /i/. The changes which led to this allophony began to occur in the latter part of the Late Middle Japanese period (1200 – 1600 AD; Frellesvig 2010), and thus the Kunrei-shiki system accurately reflects the pronunciation of /t/ and /d/ in Early Middle Japanese with no such allophony. Exceptions to this include when the spelling of a word has changed to match sound changes (e.g. Modern *azuma* ‘east’ was formerly *aduma* in Early Middle Japanese). In these cases, the Modern Japanese transcription is used (e.g. *azuma*, not *aduma*, will be used). Additionally, the phoneme /h/ has the allophones [ɸ] when preceding /ɯ/ and [ç] when preceding /i/. The digraph <sy> represents the sound [ʃ].

[4] Although often given as 2n + 1, the number of attested patterns cannot be unified into a consistent equation, as, according to Haraguchi (1979), 1 mora words show 3 patterns (2n + 1), 2 mora words show 4 patterns (2n), 3 mora words show 6 patterns (2n), and 4 mora words show 7 patterns (2n – 1). Accordingly, the characterization 2n + 1, as given by Haraguchi, with the caveat that Kyoto Japanese words show “up to” 2n + 1 patterns, seems to be the most general description. An alternative is the notation 2n(±1) used by Kubozono (2012).

systems, which may either have or lack register. Tokyo does not use registers. Modern Kyoto has two registers, one with a high beginning (“level” in Uwano’s terms), shown in (2), and one with a low beginning (“rising”), shown in (3). Words with a high beginning begin with an H tone and will stay high (hence the use of “level” in Uwano’s terms), as shown in (2a), unless there is an accent, at which point, pitch will fall, as shown in (2b-c). Note that final accent is not observed in high beginning words (Haraguchi 1999).

(2) High beginning words in Modern Kyoto

a.	SAKURA	‘cherry tree’	HHH	unaccented
b.	I’noti	‘life’	H’LL	initial accent
c.	OTO’ko	‘man’	HH’L	second accent

Unaccented words with a low beginning begin with an L tone and will stay low until the final mora, at which point, pitch will rise to H (hence “rising”), as shown in (3a). If a low beginning word is accented, as in (3b-d), the word will stay low until it reaches the accent, at which point pitch will rise to H and then fall back to L at the location of the accent. If a word has final accent with no attached particle (3c), the final mora will have a falling contour tone. However, if a word is followed by a particle such as the nominative marker *ga* as in (3d), the accentual H falls on the final mora of the lexical word, and the accentual L falls on the particle. Final accent is only observed in low beginning words (Haraguchi 1999).

(3) Low beginning words in Modern Kyoto

a.	suzuME	‘sparrow’	LLH	unaccented
b.	kaBU’to	‘helmet’	LHL	second accent
c.	osaK'i	‘the future’	LLF	final accent
d.	osaKI-ga	‘the future (Nom.)’	LLH-L	final accent

A tabular comparison of Tokyo and Modern Kyoto on these two parameters is shown in (4) below.

(4) Comparison of Tokyo and Modern Kyoto (Haraguchi 1999, Uwano 1999)

Dialect	**# of Accentual Patterns**	**# and Type of Tonal Registers**
Tokyo	n + 1	N/A
Modern Kyoto	2n + 1	2 – high beginning, low beginning

3 The Prevalence of Initial Accent

3.1 Initial Accent in Tokyo and Modern Kyoto Bimoraic and Trimoraic Nouns

Although accent may be located on any mora in an accented noun in principle as demonstrated in the previous examples, an examination of the prevalence of accent patterns reveals that possible accentual patterns are not equally prevalent. This discussion focuses on bimoraic and trimoraic nouns in the native lexicon.

I focus on nouns in the present study because the pitch accent system of nouns is usually considered to have unpredictable, underlying accent, which may fall on any mora in the word, including the initial mora, which is the most common pattern in accented nouns and the focus of the present study. Other lexical categories, such as verbs and adjectives, have a more predictable accent system, with accent falling on the antepenultimate or penultimate syllable (depending on the type of suffixes attached to these words) if present at all (Haraguchi 1999). Nouns thus provide a more interesting accentual system which can be examined to uncover the kinds of changes that are likely to have occurred over Japanese's developmental history, which can subsequently inform our understanding of how and why the proposed changes occurred and why the initial accent pattern is the most prevalent among accented nouns.

I focus on the native lexicon for three reasons: 1) words whose accentual patterns in Early Middle Japanese were recorded in the *Ruijumyogisho* consist primarily of native Japanese words; 2) the native lexicon is the gateway to understanding the native accentual system of Japanese; an understanding of this system and its development from majority non-initial accent in Early Middle Japanese accented nouns to majority initial accent in Modern Kyoto accented nouns can provide a window into both how the native system operated in the past and operates in the present; and 3) the overwhelming majority of Sino-Japanese nouns and loanword nouns have antepenultimate accent, which is the default accentual pattern for non-native nouns, suggesting that this default has origins in the language's native accentual system; accordingly, it is natural to investigate the native accentual system in order to better understand the accentual system of non-native nouns.

Finally, the rationale behind the focus on short bimoraic and trimoraic nouns lies in the fact that as the length of words increases, so too does the amount of internal morphological structure. In Tokyo, words in the native lexicon usually vary in length from one to three moras, and longer words are usually polymorphemic (Kubozono 2006). The Modern Kyoto lexicon paints a similar picture: words in the native lexicon are usually two[5] to three moras in length, and monomorphemic words four moras or longer exist but are not common (Yoshida and Zamma 2001). Compound nouns exhibit accentual behavior different from morphologically simplex nouns, in which accent falls as close as possible before or after the boundary between its morphemes (for further details, see Kubozono 1995, 2008, Labrune 2012, and Ito and Mester 2007 on Tokyo, Haraguchi 1999 and Nakai 2002 on Modern Kyoto and references cited therein). Limiting the scope of the present study to accentual patterns in morphologically simplex nouns (primarily represented by bi- and trimoraic nouns), on the other hand, allows for the examination and evaluation of the basic accentual systems of Tokyo and Modern Kyoto Japanese without interference from the compound accent system. Furthermore, the present study focuses on trimoraic nouns in particular in order to ensure that the maximal possible number of diachronic processes are considered, particularly processes which would not have applied to the shorter, bimoraic words, such as enforcement of the prohibition against two H tones in a single word (culminativity).

Before examining accentual pattern frequencies, it is worthwhile to briefly discuss lexical strata in the Japanese lexicon, as pattern frequencies vary depending on lexical

---

[5] Unlike Tokyo, Modern Kyoto has a minimum word requirement of two moras. Thus, words that are monomoraic in Tokyo, such as *na* ‘name’ and *ki* ‘tree,’ are bimoraic in Modern Kyoto as *naa* and *kii* respectively (Haraguchi 1999, Yoshida and Zamma 2001).

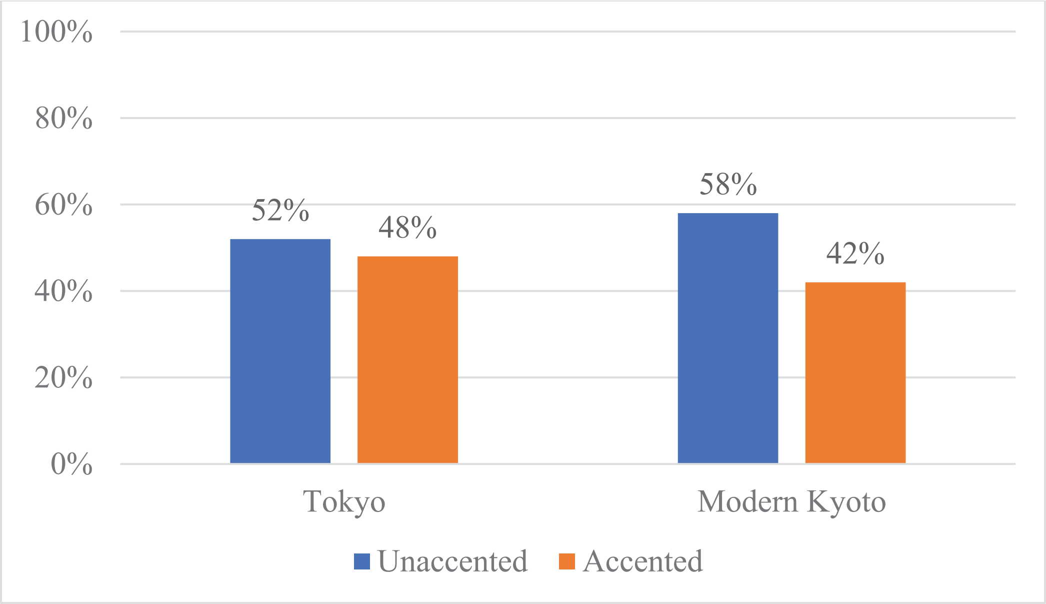
stratum. Lexical words in the Japanese lexicon can be divided into three main lexical strata by word origin: native Japanese words (*wago* ‘Japanese words’ or *Yamato-kotoba* ‘Yamato words’), Sino-Japanese words (*kango* ‘Chinese words’) and loanwords of non-Chinese origin (*gairaigo* ‘foreign words’) (Shibatani 1990). Sino-Japanese words constitute a substantial part of the Modern Japanese lexicon, representing about 60% of words in the *Genkai* dictionary (1859, as cited by Shibatani 1990) and 20% of ordinary speech (Ito and Mester 2015). Loanwords make up 1.4% of the words in the *Genkai* dictionary (Shibatani 1990). Sino-Japanese words historically originate from intensive borrowing from Chinese in the 6th, 8th, and 14th centuries (Ito and Mester 2015). Non-Chinese loanwords originate from borrowings from primarily Portuguese and Dutch from the 16th to the first half of the 19th century and primarily English from the latter half of the 19th century onward (Shibatani 1990). As will be seen, accentual pattern frequencies differ between lexical strata, although pattern frequencies are similar between the Sino-Japanese and loanword strata.

The majority of trimoraic nouns in both Tokyo and Modern Kyoto are unaccented. This is shown in the table in (5) below. The frequencies in this table account for native, Sino-Japanese, and loanword nouns combined. A bar graph representation is given in (6) below.

(5) Frequency of Accentedness in Tokyo and Modern Kyoto 3μ Nouns

Accentual Pattern	Tokyo (Kubozono 2008) n=7,937	Modern Kyoto (Sugito 1995) n=12,550
Unaccented	**52%**	**58%**
Accented	48%	42%

(6) Bar Graph Representation of (5)



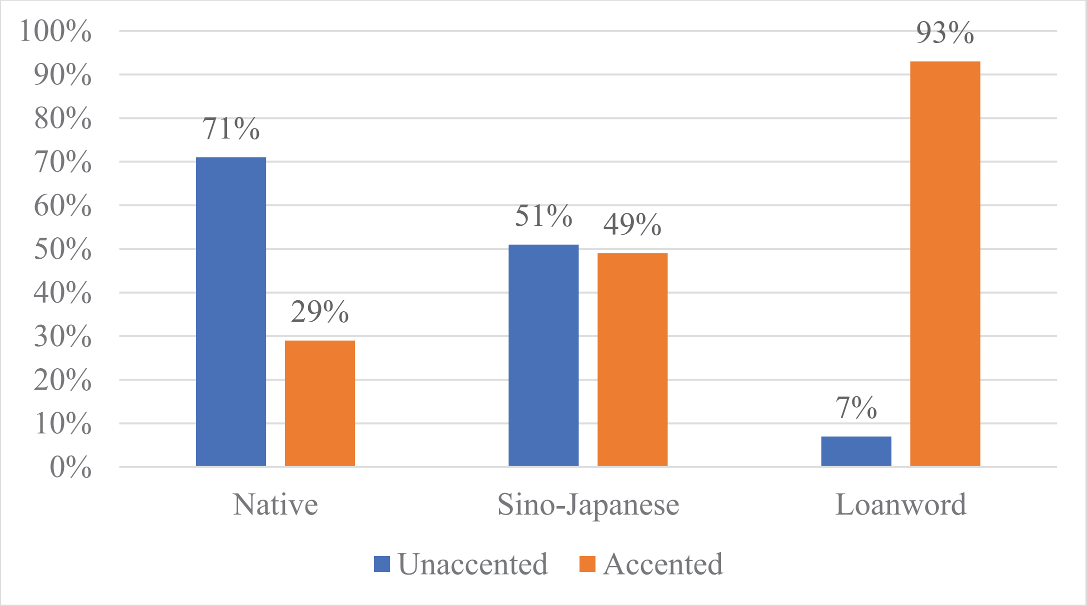
For Modern Kyoto, Sugito (1995) reports the raw numerical frequency distribution of accent patterns for 12,550 trimoraic nouns recorded from one older generation speaker of Osaka Japanese, Speaker MA (born 1932) in her electronic dictionary, the *Osaka-Tokyo Accent Voice Dictionary*. Sugito's data combine the native, Sino-Japanese, and loanword strata and cannot be separated by stratum. Using this data, I calculated the frequency of unaccented vs. accented nouns presented in the table in (5). For Tokyo, Kubozono (2008) reports accented vs. unaccented frequencies split by lexical stratum, but not overall frequencies. In order to make a comparison between Kubozono's data and the frequencies I calculated for Sugito's data, I calculated pooled frequencies using raw numerical data reported by Kubozono on 7,937 trimoraic nouns. These pooled frequencies are also presented in (5).

When accentedness data is split by lexical stratum, the following picture in (7) results for Tokyo (Kubozono 2008), showing that the ratio of unaccented words to accented words changes depending on lexical stratum. A bar graph representation is given in (8).

(7) Frequency of Accentedness in Tokyo 3μ Nouns

Accentual Pattern	Native n=2,220	Sino-Japanese n=4,939	Loanword n=778
Unaccented	71%	51%	7%
Accented	29%	49%	93%

(8) Bar Graph Representation of (7)



As (7) and (8) show, the proportion of unaccented nouns decreases from being the majority in native nouns, approximately equal in prevalence to accented nouns in

Sino-Japanese nouns, to being a small minority in loanwords. That is, unaccentedness is less prevalent in loans than in native words, and less prevalent in newer loans than in older loans.

On the other hand, Sugito (1995)'s dictionary was not designed to be able to search frequencies by lexical stratum, so it is not possible to provide data for Modern Kyoto from Sugito comparable to the data in (7). However, other literature shows a similar trend, with much lower rates of unaccentedness in non-Sino-Japanese loanwords when compared to native words. Yoshida and Zamma (2001) surveyed 115 morphologically simplex native three mora nouns from Modern Kyoto and report raw numerical frequencies for each accent pattern. Calculating frequencies from these raw data, the following picture results for unaccented vs. accented frequencies in native nouns, shown in (9), rounded to the nearest whole number.

(9) Frequency of Accentedness in Morphologically Simplex Modern Kyoto 3μ Native Nouns (n=115; Yoshida and Zamma 2001)

Accentual Pattern	Frequency
Accented	**54%**
Unaccented	46%

Interestingly, the data considered by Yoshida and Zamma show a slight preference for accented over unaccented nouns. However, because they surveyed only 115 words, an examination of a larger sample, on a scale such as those considered by Kubozono (2006, 2008) and by Sugito (1995), may reveal a preference for unaccented words.

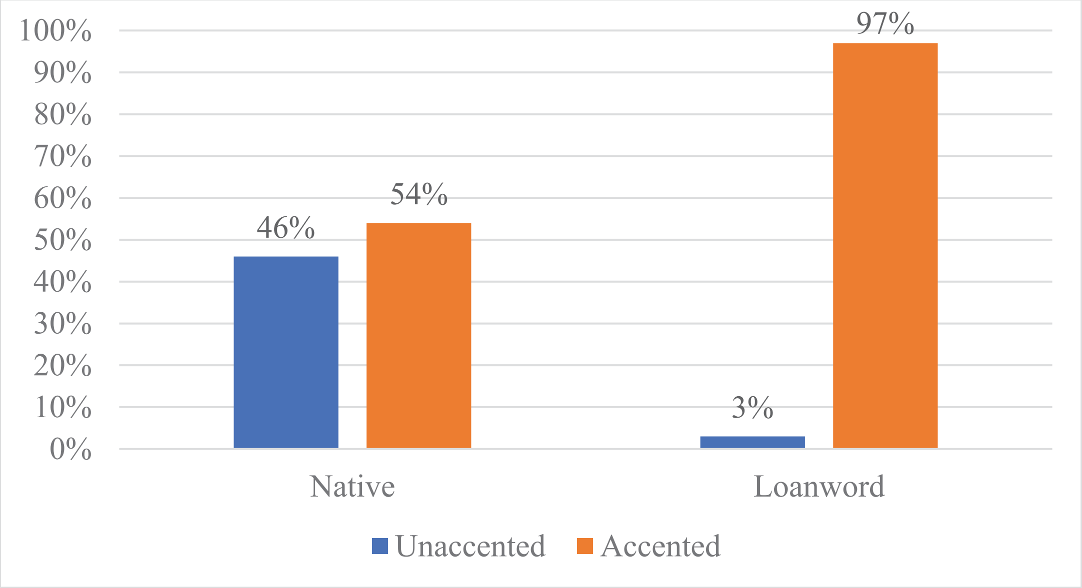
Non-Sino-Japanese trimoraic loanwords are overwhelmingly accented on the other hand. Data from Tanaka (2018) show the following split in unaccented vs. accented nouns in (10), rounded to the nearest whole number.

(10) Frequency of Accentedness in Modern Kyoto 3μ Loanword Nouns (n=91; Tanaka 2018)

Accentual Pattern	Frequency
Accented	**97%**
Unaccented	3%

A bar graph representation of (9) and (10) is given in (11) below.

(11) Bar Graph Representation of (9) and (10)

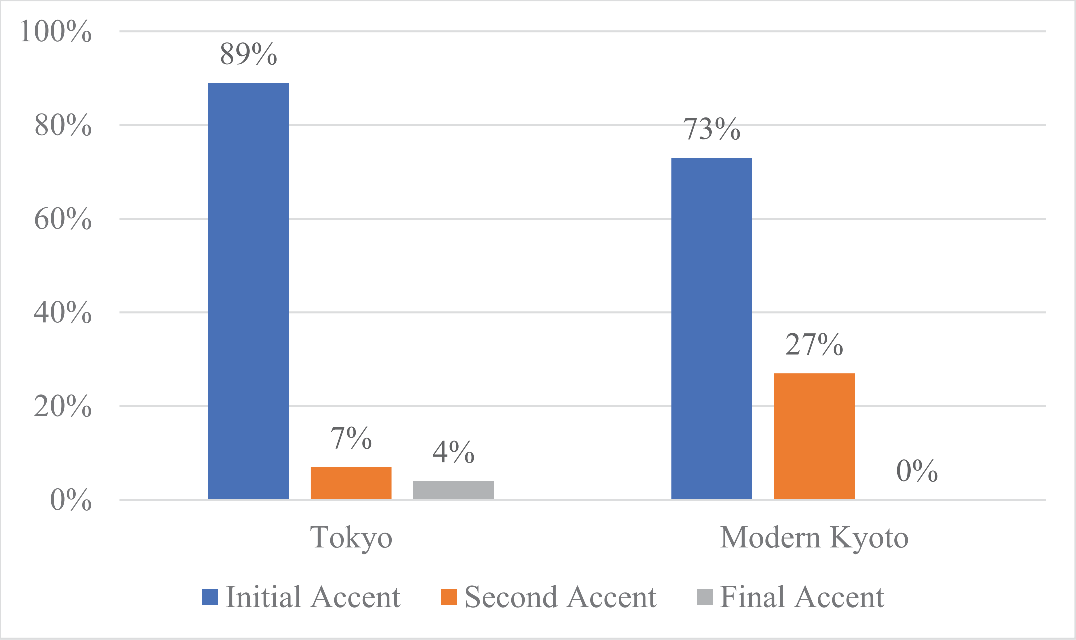


Focusing our view now on only accented nouns, we find that the majority of accented trimoraic nouns have initial accent. Frequencies are shown in the table in (12) below, again combining native, Sino-Japanese, and loanword frequencies, rounded to the nearest whole number. A bar graph representation is given in (13).

(12) Frequency of Accent Patterns in Tokyo and Modern Kyoto Accented 3μ Nouns

Accentual Pattern	Tokyo (Kubozono 2008) n=3,783	Modern Kyoto (Sugito 1995) n=5,216
Initial Accent	**89%**	**73%**
Second Accent	7%	27%
Final Accent	4%	N/A

(13) Bar Graph Representation of (12)



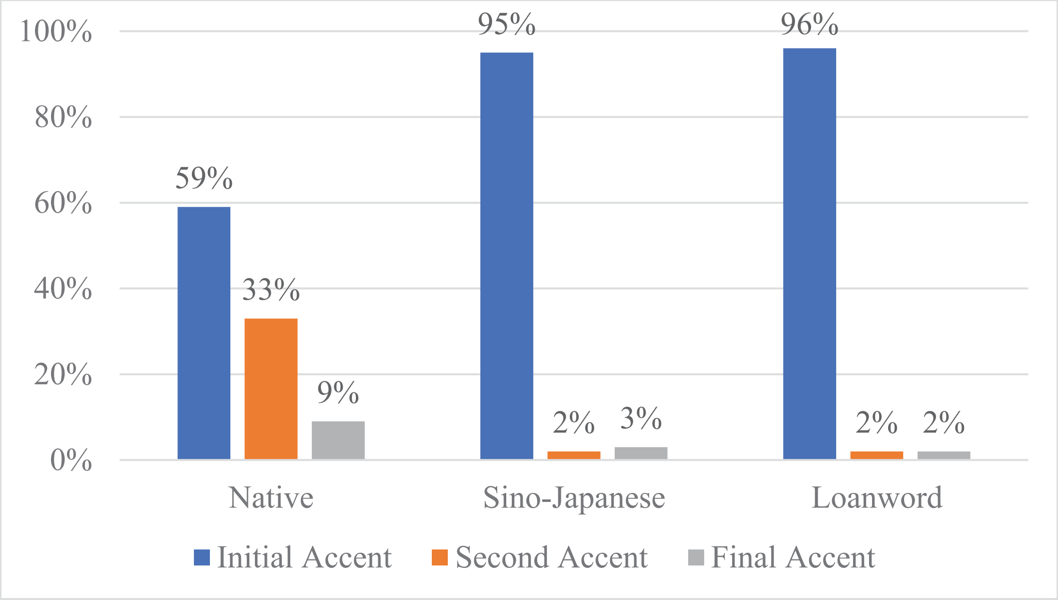
As with the frequencies in (5), I calculated Modern Kyoto frequencies using the raw data reported by Sugito (1995) on 5,216 accented nouns recorded from Speaker MA. Kubozono (2008)'s Tokyo data is based on 3,783 nouns.

Combining the frequencies of native, Sino-Japanese, and loanword trimoraic nouns will necessarily result in skewed initial accent frequencies because frequencies vary by lexical stratum as shown in (7) through (11) above, so examining the frequencies for each lexical stratum in both dialects is also necessary. Table (14) below shows accentual pattern frequencies in accented nouns split by lexical stratum in Tokyo as reported by Kubozono (2008). A bar graph representation is given in (15).

(14) Frequency of Accent Patterns in Tokyo Accented 3μ Nouns by Lexical Stratum (n=3,783; Kubozono 2008)

Accentual Pattern	Native n=634	Sino-Japanese n=2,427	Loanword n=722
Initial Accent	**59%**	**95%**	**96%**
Second Accent	33%	2%	2%
Final Accent	9%	3%	2%

(15) Bar Graph Representation of (14)



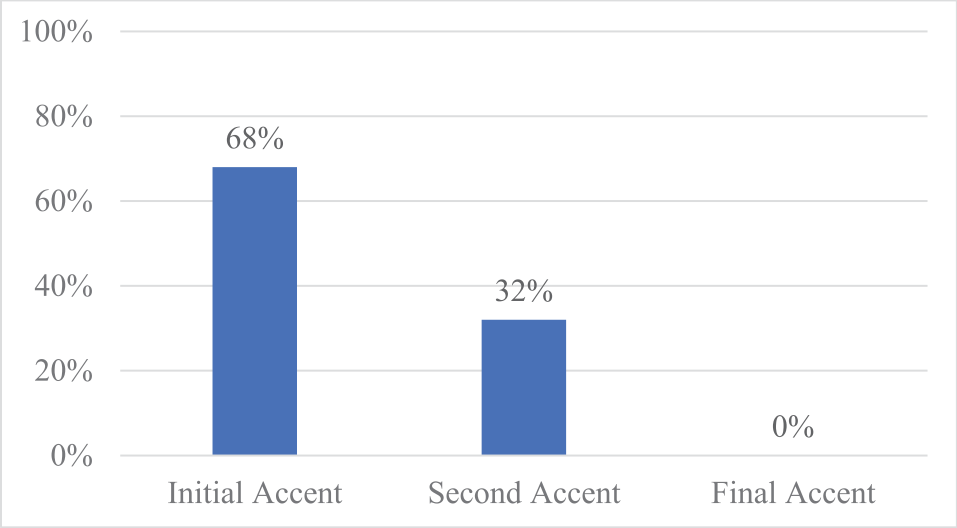
Kubozono (2008) considered 634 native words, 2,427 Sino-Japanese words, and 722 loanwords. As table (14) shows, separating accentual pattern frequencies by lexical stratum still reveals that initial accent is the majority pattern at 59%. However, a significant minority of native trimoraic nouns have accent on the second syllable. Kubozono (2006) explains that a substantial amount of these second accented trimoraic nouns are in fact short compounds, such as *haNA'-ya* 'flower shop.' When frequencies are corrected to consider only morphologically simplex trimoraic nouns, Kubozono (2006) reports that the prevalence of initial accent jumps up to about 70%.

As discussed above, a comparable examination of Sugito (1995) is not possible. However, examinations of morphologically simplex three mora native nouns and three mora loanwords from other literature shows that both strata trend approximately equally toward initial accent. Yoshida and Zamma's survey of 115 morphologically simplex three mora nouns included 62 accented nouns. The table in (16) shows frequencies for each accentual pattern among the 62 accented nouns included in the survey, and a bar graph representation is given in (17).

(16) Frequency of Accent Patterns in Accented Morphologically Simplex Modern Kyoto 3μ Nouns (n=62; Yoshida and Zamma 2001)

Accentual Pattern	Frequency
Initial Accent	**68%**
Second Accent	32%
Final Accent	N/A

(17) Bar Graph Representation of (16)



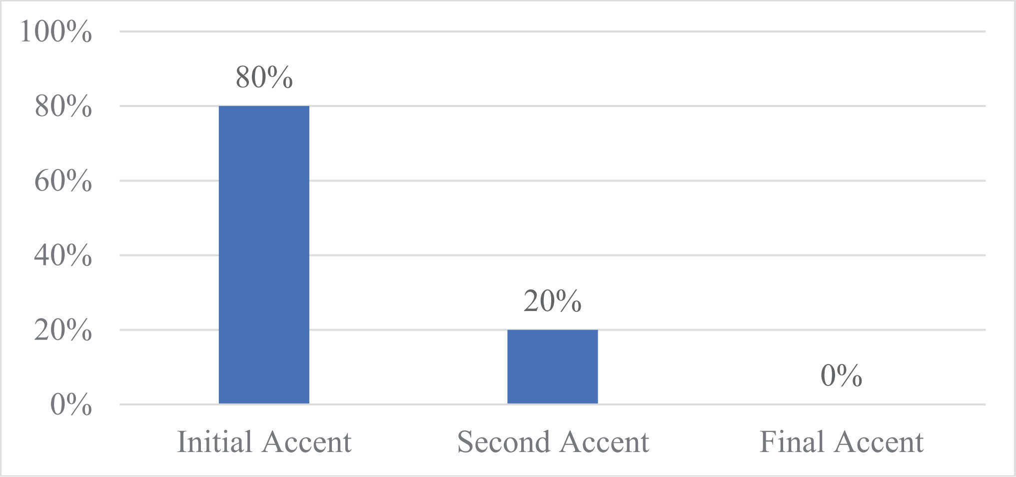
The frequency reported for initial accent in accented words is as expected, with a majority exhibiting initial accent.

Tanaka (2018)'s survey of 91 three mora loanword nouns in Modern Kyoto included 88 accented nouns, whose accentual patterns are summarized in table (18). A bar graph representation is given in (19).

(18) Frequency of Accent Patterns in Modern Kyoto 3μ Accented Loanword Nouns (n=88; Tanaka 2018)

Accentual Pattern	Frequency
Initial Accent	**80%**
Second Accent	20%
Final Accent	N/A

(19) Bar Graph Representation of (18)



Because of the frequency of initial accent in accented words, Kubozono (2006, 2008) and Yoshida and Zamma (2001) interpret the results of their investigations as supporting the idea that antepenultimate accent (or initial accent in words three moras or shorter) is truly a default, even in native nouns, which are usually considered to have unpredictable underlyingly-specified accent, as discussed in the introduction. Based on this finding, Kubozono (2008) proposes that the accent system of Tokyo is closer to that of an N-pattern system with two patterns than a multi-pattern system. Kubozono proposes that the two patterns of Tokyo are accented vs. unaccented, where accented nouns receive default antepenultimate accent. The remainder of the nominal lexicon, consisting only of those nouns which are neither unaccented nor initial accented, then has lexically specified accent location. Regardless of whether this interpretation of the data should be regarded as true, it is remarkable that the initial accent pattern is so prevalent, even among morphologically simplex nouns in the native stratum.

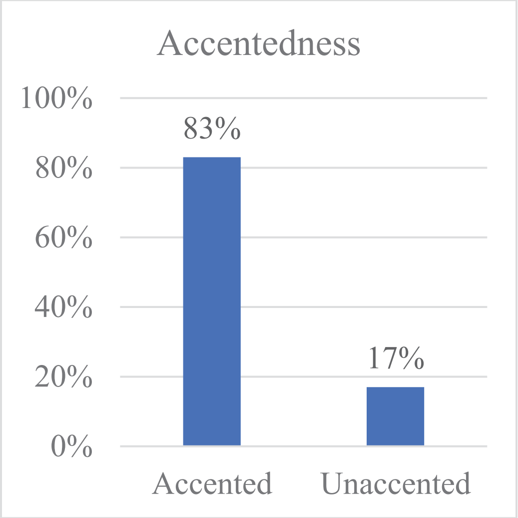
Turning to bimoraic nouns, the same pattern arises in both Tokyo and Modern Kyoto, with initial accent constituting the clear majority – 66% of bimoraic nouns in Tokyo and 65% of bimoraic nouns in Modern Kyoto. However, unlike trimoraic nouns, whose most common pattern as a whole is unaccented, with initial accent being the most frequent among accented nouns, initial accent is the most common pattern in bimoraic nouns regardless of accentedness. The percentages below, rounded to the nearest whole number, are calculated from raw frequency data provided in Sugito (1995). Tokyo frequencies are given in (20) with a bar graph representation in (21), and Modern Kyoto frequencies for nouns produced by Speaker MA are given in (22), with a bar graph representation given in (23). As discussed above, it is not possible to separate frequencies by lexical stratum or morphological complexity with Sugito's dictionary.

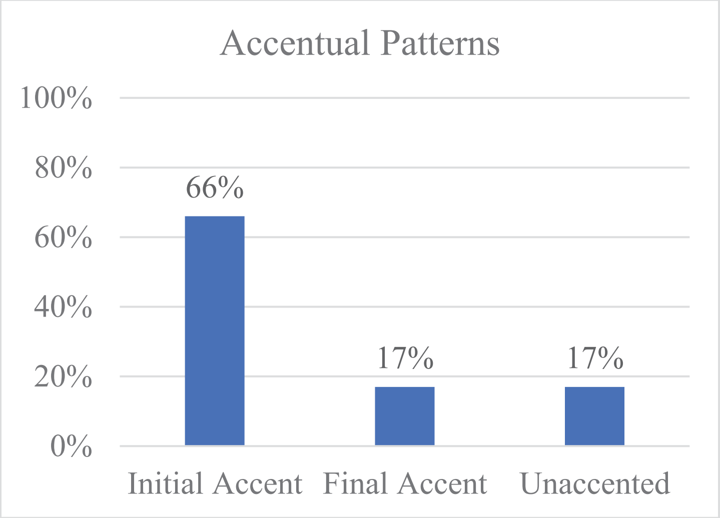
(20) Frequency of Accent Patterns in Tokyo 2μ Nouns (n=3,019; Sugito 1995)

Accentedness	Frequency
Accented	**83%**
Unaccented	17%

Accentual Patterns	Frequency
Initial Accent	**66%**
Final Accent	17%
Unaccented	17%

(21) Bar Graph Representations of (20)



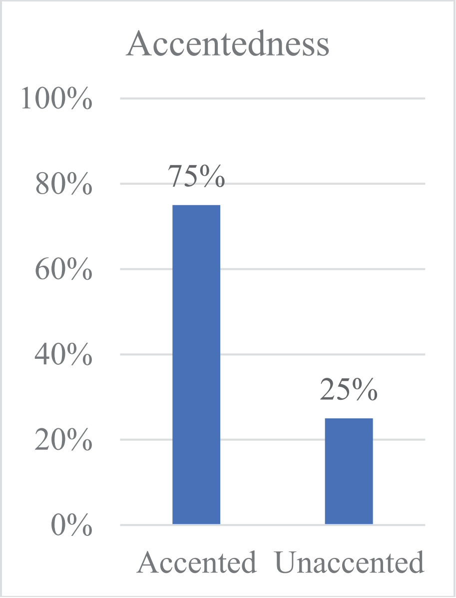


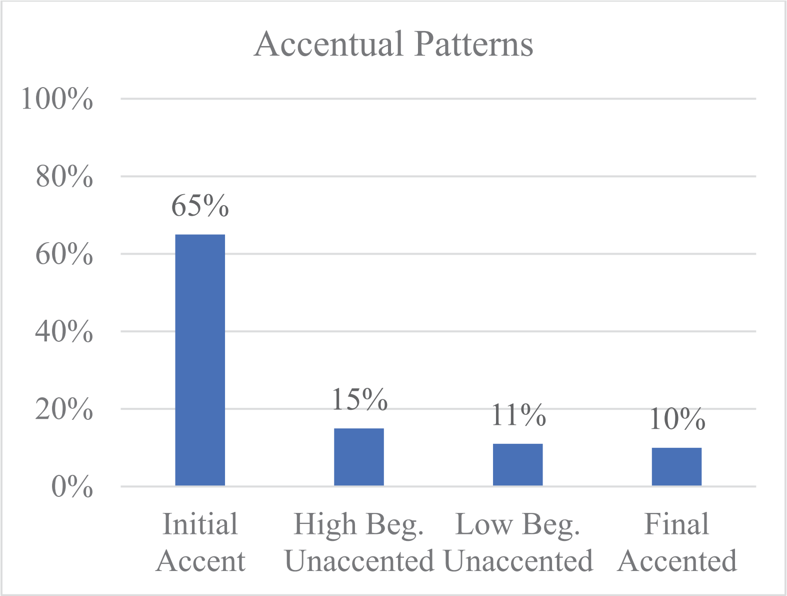
(22) Frequency of Accent Patterns in Modern Kyoto 2µ Nouns (n=3,019; Sugito 1995)

Accentedness	Frequency
Accented	**75%**
Unaccented	25%

Accentual Patterns	Frequency
Initial Accent	**65%**
High Beginning Unaccented	15%
Low Beginning Unaccented	11%
Final Accent	10%

(23) Bar Graph Representations of (22)





Considering only accented nouns, initial accent makes up 80% of accented bimoraic nouns in Tokyo (n=2510) and 86% of accented bimoraic nouns in Kyoto (n=2259).

Despite the prevalence of initial accent as the second most common pattern in trimoraic nouns after unaccentedness and the most common pattern in bimoraic nouns, nouns in Early Middle Japanese, the ancestor of Modern Kyoto, present a different picture. I will focus on the development from Early Middle Japanese to Modern Kyoto for the remainder of the paper.

3.2 Initial Accent in Early Middle Japanese Bimoraic and Trimoraic Nouns

Sugito (1995)'s dictionary also includes accentual information for 538 bimoraic nouns and 420 trimoraic nouns from Early Middle Japanese (EMJ)[6], whose accentual patterns are recorded in a 12th century Japanese-Chinese dictionary called the *Ruijumyogisho*. In contrast to the majority prevalence of initial accent in Tokyo and Modern Kyoto, an examination of accent pattern frequencies in this sample of the *Ruijumyogisho* reveals that initial accent in EMJ was relatively uncommon.

I calculated accentual pattern frequencies from the raw numerical data reported by Sugito (1995) on the EMJ accent of these 538 bimoraic nouns and 420 trimoraic nouns. The results of these calculations are presented in the tables in (24-25) and (27-28) below. The frequencies in each table are presented in descending order from most frequent to least frequent. The tables in (24) and (27) present unaccented patterns in the sense of Haraguchi (1979)'s description of Old Kyoto, in which words which do not have pitch changes are considered unaccented, while words which do have a pitch

---

[6] In total, Sugito (1995) includes 2365 EMJ words, of which 1265 are nouns. The remaining 307 nouns whose data was not included here consist of nouns four moras in length or longer.

change are considered accented. Thus, LLH, which is considered unaccented in Modern Kyoto, is considered accented in Old Kyoto. This distinction is observed here to illustrate the relatively low prevalence of initial accent among both accented nouns alone and all nouns together as well as to illustrate the relative stability of the unaccented majority in trimoraic nouns and the shift from majority unaccented to majority accented in bimoraic nouns across the development of Modern Kyoto. The tables in (25) and (28) present the accented patterns. An example word for each pattern is also provided.

First, let us consider trimoraic nouns, presented in (24) and (25). Unaccented nouns are given in (24), while accented nouns, which have a pitch change, are given in (25). A bar graph representation of the accented noun patterns in (25) is given in (26). The pattern of interest, HLL, is bolded and enclosed with bolded borders in (25) and enclosed in a box in (26).

(24) Frequency of Accent Patterns in 3μ Unaccented Nouns (n=259; 61.7% of 3μ Nouns) in the *Ruijumyogisho* (Sugito 1995)

Accentual Pattern	Frequency	% of 3μ Unaccented Nouns	% of all 3μ Nouns	Example
HHH	133	51.4%[7]	31.7%	*katati* ‘shape’
LLL	126	48.6%	30.0%	*otoko* ‘man’
Total	259	100.0%	61.7%	

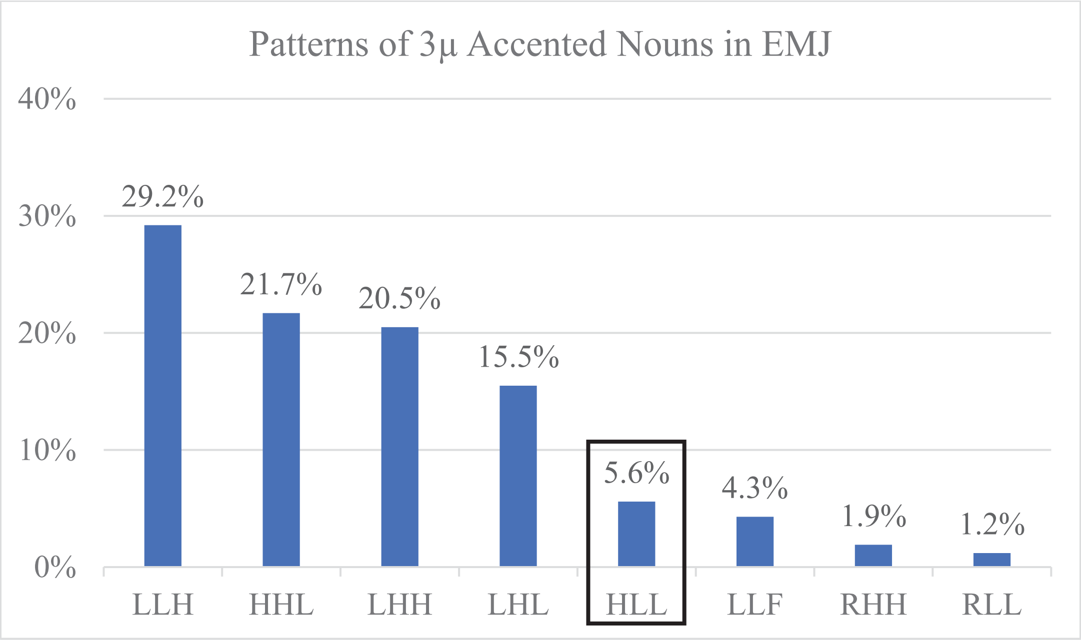
---

[7] Percentages are rounded to the nearest tenth. As a result, the total percentage of 3μ accented nouns (99.9%) and the combined total percentages of all 3μ nouns (61.7% + 38.4% = 100.1%) do not equal 100%.

(25) Frequency of Accent Patterns in 3μ Accented Nouns (n=161; 38.3% of 3μ Nouns) in the *Ruijumyogisho* (Sugito 1995)

Accentual Pattern	Frequency	% of 3μ Accented Nouns	% of all 3μ Nouns	Example
LLH	47	29.2%	11.2%	*abura* 'oil'
HHL	35	21.7%	8.3%	*tobira* 'door'
LHH	33	20.5%	7.9%	*suzume* 'sparrow'
LHL	25	15.5%	6.0%	*hatake* 'cultivated land'
HLL	**9**	**5.6%**	**2.1%**	***tikara* 'power'**
LLF	7	4.3%	1.7%	*yorozu* 'ten thousand'
RHH	3	1.9%	0.7%	*hisui* 'jade'
RLL	2	1.2%	0.5%	*zisyaku* 'magnet'
Total	161	99.9%	38.4%	

(26) Bar Graph Representations of (25)



As the table in (25) shows, initial accent (HLL) is quite rare at 5.6% of accented 3μ nouns, or 2.1% of all 3μ nouns, appearing to pattern more closely with words with an accentual pattern featuring a contour tone than with words featuring only level tones. This is in stark contrast to the 68% prevalence of initial accent in Modern Kyoto accented trimoraic nouns. In addition, while the nine words of the HLL pattern and the seven words of the LLF pattern are all native Japanese words, all five of the RHH and RLL words are Sino-Japanese, making HLL the second rarest accent pattern among native nouns.

Note that 61.7% of the nouns considered here are unaccented (HHH and LLL) as shown in (24). This is very close to the 58% 20th century frequency of unaccented patterns in Modern Kyoto as reported in table (5). It appears, then, that while the frequency of unaccented patterns appears to have been roughly stable since the 12th century, the prevalence of initial accented nouns has increased significantly since then.

Next, let us consider bimoraic nouns, presented in (27) and (28). A bar graph representation of the accented nouns in (28) is given in (29). The pattern of interest, HL, is bolded and enclosed in bolded borders in (28) and enclosed in a box in (29).

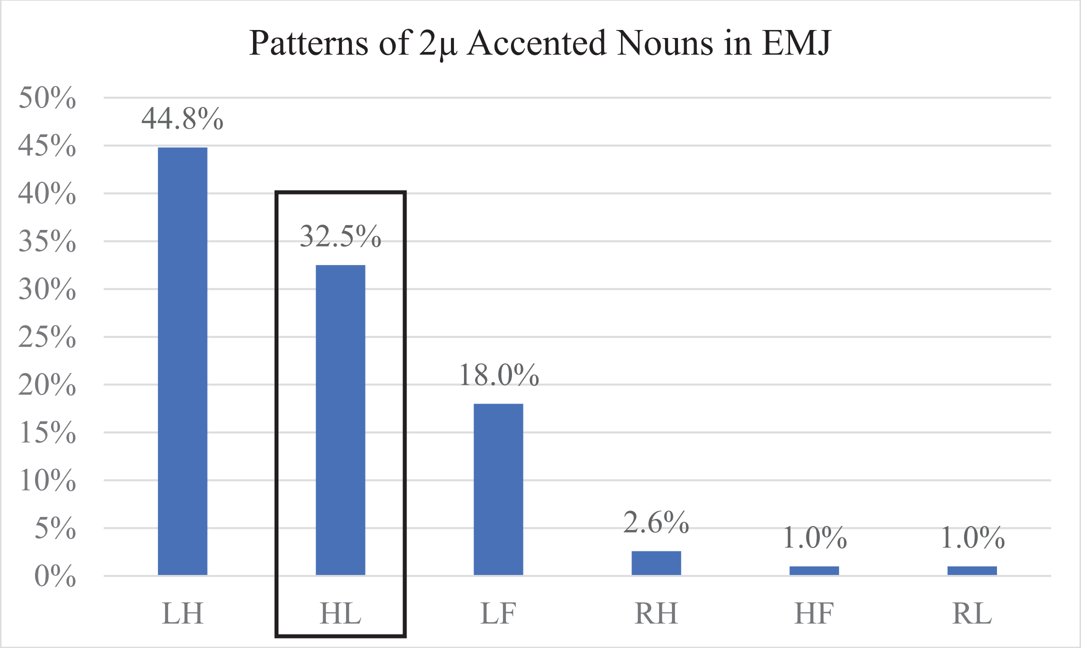
(27) Frequency of Accent Patterns in 2μ Unaccented Nouns (n=348; 64.2% of 2μ nouns) in the *Ruijumyogisho* (Sugito 1995)

Accentual Pattern	Frequency	% of 2μ Unaccented Nouns	% of all 2μ Nouns	Example
LL	183	52.6%	33.8%	*kami* 'spirit, god'
HH	165	47.4%	30.4%	*hana* 'nose'
Total	348	100%	64.2%	

(28) Frequency of Accent Patterns in 2μ Accented Nouns (n=194; 35.8% of 2μ nouns) in the *Ruijumyogisho* (Sugito 1995)

Accentual Pattern	Frequency	% of 2μ Accented Nouns	% of all 2μ Nouns	Example
LH	87	44.8%	16.1%	*wara* 'straw'
HL	**63**	**32.5%**	**11.6%**	***hito* 'person'**
LF	35	18.0%	6.5%	*ame* 'rain'
RH	5	2.6%	0.9%	*yuri* 'lily'
HF	2	1.0%	0.4%	*mizo* 'ditch'
RL	2	1.0%	0.4%	*hagi* 'leg'
Total	194	99.9%	35.9%	

(29) Bar Graph Representation of (28)



As shown in table (28), the initial accented pattern, HL, makes up only 11.6% of all 2μ nouns and 32.5% of 2μ accented nouns. Unlike the patterns with contour tones in three mora nouns, the majority of words with contour tones in two mora nouns are native nouns. Only four of the 44 bimoraic nouns with contour tones are Sino-

Japanese words, and all four belong to the RH accent class. Again, the difference in frequency of initial accent in bimoraic nouns between EMJ and Modern Kyoto is as striking as the difference observed in trimoraic nouns: 11.6% of all bimoraic nouns are HL in EMJ, and 64.5% of all bimoraic nouns are HL in Modern Kyoto. Considering only accented nouns, 32.5% of all accented bimoraic nouns are HL in EMJ, and 85% of all accented bimoraic nouns are HL in Modern Kyoto.

4 Accentual Patterns from Early Middle Japanese to Modern Kyoto

4.1 The *Ruijumyogisho* and the Interpretation of its Tone Markings

The earliest records of a Japanese accentual system can be found in the *Ruijumyogisho* (Frellesvig 2010). The *Ruijumyogisho* is a 12th century Chinese-Japanese dictionary in which scholars recorded the accentual system of the dialect of Early Middle Japanese (EMJ; also called Old Kyoto Japanese) spoken in Kyoto in the late Heian era in the 12th century, which is regarded as the ancestor dialect to Modern Kyoto (Haraguchi 1979, Frellesvig 2010). There are no records for Tokyo accent before the modern period (Frellesvig 2010). Tones were recorded in the *Ruijumyogisho* using a notation method Chinese scholars used to record Chinese tones in dictionaries. In the system used in the *Ruijumyogisho*, a series of "tone dots" were placed next to syllabic *kana* characters to mark tone. Although there were four tone dots, the most prevalent tone dots were the 上 "rising" tone dot and the 平

“even” tone dot[8], which are traditionally understood to have referred to the H and L tones (Shimabukuro 2007, Frellesvig 2010). The rising tone dot was represented with a dot placed to the upper left of a character, and the even tone dot was represented with a dot placed to the lower left of a character. The placement of the rising and even tone dots is demonstrated in the schematics in (30) and (31) below with the *kana* character カ for the syllable *ka*.

(30) “Rising” tone dot, indicating H tone: *ká*

(31) “Even” tone dot, indicating L tone: *kà*

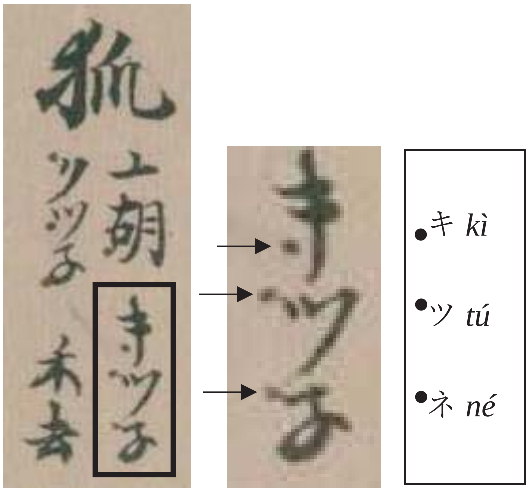
The following examples demonstrate actual usage of the tone dots in the *Ruijumyogisho*. (32) depicts the entry 狐 *kìtúné* ‘fox’ (accentual pattern LHH), and (33) depicts the entry for 疾 *yàmâì* ‘illness’ (accentual pattern LHL). In each entry, the Japanese readings of each character written in the *katakana* syllabary with tone markings are enclosed in boxes. An enlarged image of the readings is provided to the right of each entry, with arrows pointing to the tone dots. Finally, a transcription is

---

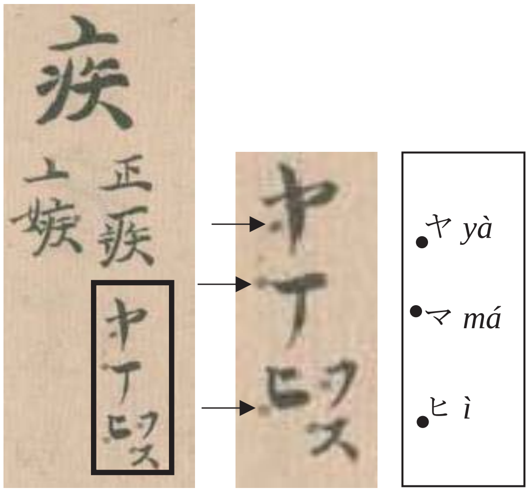
[8] The other two tone dots were the 東 “east” tone dot, which was represented by a dot placed around the midline to the left of a character, and the 去 “departing” tone dot, which was represented by a dot placed to the upper right of a character. They are used to mark the falling and rising tones respectively, but the east tone dot was not used in all copies of the *Ruijumyogisho*, and neither were otherwise used regularly or consistently (Ramsey 1979, Frellesvig 2010).

provided to the right of the enlarged image in both modern *katakana* with tone dot markings and their romanizations with tones indicated with acute and grave diacritics. Entries are taken from a digital copy of the *Ruijumyogisho* hosted by the National Diet Library, Japan's Digital Collections.

(32) *Ruijumyogisho* entry for *kìtúné* 'fox'



(33) *Ruijumyogisho* entry for *yàmáì* 'illness'



Although the rising tone dot is traditionally interpreted as H, and the even tone dot is traditionally interpreted as L, not all scholars agree (e.g. Ramsey 1979, de Boer 2017), interpreting them as L and H respectively instead, citing lack of evidence of the Middle Chinese values of the tone dots, lack of evidence of how the Japanese used tone dots for Japanese, and modern Tokyo and Kyoto Japanese accent patterns which are claimed to be "better" explained with the reversed interpretation. However, Mei (1970) concludes that the rising and even tone dots do indeed represent a high tone and a low tone in Middle Chinese, respectively, based primarily on investigation of a 9th century text, *Hsi-t'an tsang*, by the Japanese Buddhist monk Annen which provides a description of the meanings of the tone dots. In this document, Annen writes that the rising tone is "level and high" and that the even tone is "level and low." Additionally, Mei considers evidence from the Japanese tradition of *bombai*, Sanskrit psalmody which was transliterated into Chinese with the appropriate tones. According to Mei, the *bombai* tradition of the Shingon sect of Japanese Buddhism follows rules for pronouncing the tones which agree with Annen's description of the rising and even tones as high and low respectively and provide evidence for this interpretation by the writers of the *Ruijumyogisho*. Building on the philological evidence discussed by Mei, Vovin (1990) rejects Ramsey's hypothesis on the basis of evidence from early Korean loanwords in Old Japanese and late first millennium Japanese loanwords in Proto-Ainu which he argues confirm the rising tone = H and even tone = L interpretation. The discussion presented here is based on the traditional understanding of the *Ruijumyogisho* tone dots with the rising tone signifying H and the even tone signifying L.

4.2 Overview of the Development from EMJ to Modern Kyoto

The accentual system of the 12th century Kyoto dialect of EMJ recorded in the *Ruijumyogisho* is, as mentioned previously, commonly regarded as the ancestral system of Modern Kyoto. It is also traditionally regarded as the ancestral system of all Japanese dialects, though recent scholarship (e.g. Matsumori 1999, 2008, Shimabukuro 2007, and Frellesvig 2010) argues that this is in fact not the case. Matsumori (1999, 2008) argues that EMJ cannot be the ancestral system of all Japanese dialects because accent data from several dialects of the Shikoku area are problematic for the traditional assumption. On this basis, Matsumori proposes a reconstruction of a stage earlier than EMJ which she calls Proto-Mainland Japanese (PMJ)[9] from which the Shikoku dialects and the EMJ system would have split. Matsumori says nothing further about whether either, both, or none of the Tokyo and Modern Kyoto branches descended from EMJ or from Proto-Mainland Japanese. Frellesvig (2010) argues that EMJ is the predecessor to Modern Kyoto, due to the similarity between the EMJ and Modern Kyoto forms, but not other dialects, which he argues split away from EMJ at an earlier stage. Shimabukuro (2007) argues that EMJ developed on its own developmental pathway from Proto-Mainland Japanese,

---

[9] As opposed to Proto-Ryukyuan, the proto-language of the Ryukyuan branch of the Japonic languages spoken in the Ryukyu Islands. Proto-Mainland Japanese and Proto-Ryukyuan are posited to be the daughter languages of Proto-Japonic (see Shimabukuro 2007 and references cited therein for further details).

while Modern Kyoto, Tokyo, and other accentual systems such as Kagoshima developed on separate pathways from PMJ. Among Shimabukuro's arguments against the assumption that Modern Kyoto descends directly from EMJ, Shimabukuro argues that work by Kindaichi (1960) had presented a development pathway in which words in two accentual classes had the form HLL(L) in the 14th century but had failed to merge, with one accentual class becoming HLL(L) in Modern Kyoto and the other accentual class becoming HHL(L) in Modern Kyoto. Calling this a discrepancy, Shimabukuro says that there is no reason the two accentual classes should not have completely merged in the 14th century, yielding only HLL(L) in Modern Kyoto, and argues that this provides evidence against the assumption that Modern Kyoto is descended directly from EMJ. It should be noted that the examples Shimabukuro gives for these two classes seem to differ in their morphological complexity. For 14th century HLL(L) → Modern Kyoto HLL(L), he gives the morphologically simplex word *inoti* 'life,' but for 14th century HLL(L) → Modern Kyoto HHL(L), he gives the transparently morphologically complex word *hitoe* 'one layer,' which is made up of *hito* 'one' and the suffix *e* 'layer,' which is used when counting layers. However, in the case of *hitoe,* due to the overall non-productivity of the suffix *e* (it can only be used with the native Japanese numbers 1-10), it may be that *hitoe* is a lexicalized compound and thus not subject to a compound accent rule. Additionally, Nakai (2001) writes that while some 2 mora + 1 mora compounds (e.g. *asi-ba* 'scaffold' lit. 'foot-place') which were originally HHL became HLL, other 2 mora + 1 mora compounds remained HHL (e.g. *bara-ka* 'rose family').

While I follow this scholarship in assuming that Tokyo developed from PMJ and not EMJ, I pursue the traditional viewpoint that Modern Kyoto developed directly from EMJ, and this assumption will form the basis of the core argument of this analysis.

Investigations into the development of Japanese accentual systems is traditionally based on a system of classification of words into word classes proposed by Kindaichi and Wada (1955) and expanded in Kindaichi (1974, both as cited in Matsumori 2001) to account for the tonal patterns found in the *Ruijumyogisho* and to track correspondences between the *Ruijumyogisho* and the modern dialects. The Kindaichi system is very well established in the field of Japanese accentology, and much scholarship on diachronic accentology proceeds with reference to it to the point that when new classes are discovered, if they are similar enough to existing classes in the Kindaichi system, they are regarded as subclasses to existing classes (Frellesvig 2010). For trimoraic nouns, Kindaichi proposed that there were seven classes (Matsumori 2001). These classes are listed in (34) below. They are usually numbered according to the Kindaichi numbering system, given in the second column of (34). However, for readability, I refer to these classes as Classes 1 to 7 in the first column of (34) and for the remainder of the paper.[10] The letters a and b in Class 5 represent two subclasses of Class 5.[11] Examples of each class are provided as well.

---

[10] The Kindaichi numbering system consists of the mora count, a period (.), and the class number, so that class 2.1 is the first class among two mora nouns, class 3.2 is the second class among three mora nouns, and so on.

[11] According to Nakai (2001), few words belong to Class 5b. According to Frellesvig (2010), Class 5 is separated into 5a and 5b because some words that are usually marked LLH are marked as LLF in some sources. He also notes that Classes 5a and 5b completely merged soon after EMJ, making it difficult to

(34) Kindaichi’s seven classes for trimoraic nouns in the *Ruijumyogisho* (Frellesvig 2010)

Class 1	3.1	HHH	*katati* ‘shape’
Class 2	3.2	HHL	*azuki* ‘red bean’
Class 3	3.3	HLL	*tikara* ‘strength’
Class 4	3.4	LLL	*otoko* ‘man’
Class 5a	3.5a	LLH	*abura* ‘oil’
Class 5b	3.5b	LLF	*tamaki* ‘elbow guard’
Class 6	3.6	LHH	*usagi* ‘rabbit’
Class 7	3.7	LHL	*kabuto* ‘helmet’

Correspondences can be established between these classes and the accentual patterns seen in trimoraic nouns in Modern Kyoto. These correspondences are shown in (35) and (36) below. Correspondences from EMJ to Modern Kyoto initial accent (HLL) are given in (35), while correspondences to accentual patterns without initial accent are given in (36). The correspondences in (35) and (36, Classes 1, 6, and 7) are taken from Frellesvig (2010), the correspondences in (36, Classes 2, 5a, and 5) are taken from Sugito (1995), and the Modern Kyoto forms in both (35) and (36) are taken from Sugito (1995). Acute diacritics represent H, grave diacritics represent L, and circumflex diacritics represent F.

---

establish which words originally belonged to Class 5b. It is unclear whether the merger happened before Classes 5a and 5b became HLL (in which case the two subclasses likely would have merged into LLH before eventually becoming HLL) or as a result of both becoming HLL. Although he does not discuss this explicitly, Nakai (2001)’s presentation of the relevant changes suggest that the two subclasses merged because of their changing to HLL during the Nanbokucho era (14th century).

(35) EMJ to Modern Kyoto initial accent correspondences in 3μ Nouns

Class	EMJ	EMJ Examples	Modern Kyoto	Modern Kyoto Examples
Class 2	HHL	*ázúkì*	HLL	*ázùkì*
Class 3	HLL	*tíkàrà*	HLL	*tíkàrà*
Class 4	LLL	*òtòkò*	HLL ~ HHL	*ótòkò* (younger speakers) ~ *ótókò* (older speakers)
Class 5a	LLH	*àbùrá*	HLL	*ábùrà*
Class 5b	LLF	*tàmàkî*	HLL	*támàkì*

(36) EMJ to Modern Kyoto non-initial accent correspondences in 3μ Nouns

Class	EMJ	EMJ Examples	Modern Kyoto	Modern Kyoto Examples
Class 1	HHH	*kátátí*	HHH	*kátátí*
Class 2	HHL	*kásíwà* 'oak'	LHL	*kàsíwà*
Class 5a	LLH	*àgùrá* 'sitting cross-legged'	HHH	*ágúrá*
Class 5b	LLF	*nànàê* 'seven layers'	HHL	*nánáè*
Class 6	LHH	*ùságí*	LLH	*ùsàgí*
Class 7	LHL	*kàbútò*	LHL	*kàbútò*

As the correspondences in (35) show, five accent classes merged into initial accent (HLL). There is no property on which all five classes can be unified, but Class 2 HHL and Class 3 HLL both begin with H and have the existence of an HL pitch fall in common, while Class 4 LLL, Class 5a LLH, and Class 5b LLF all begin with two L toned moras. The fact that the change from HHL, LLL, LLH, and LLF to HLL cannot be unified into a single process suggests that several changes must have occurred in the development of initial accent. I posit that these processes were: 1) change of the first tone in Classes 4, 5a, and 5b to H, 2) loss of the final H and F tones in Classes 5a and 5b by prohibition of multiple separate H tones, and 3) reduction of multiply linked H tones in Class 2. In addition, I posit that these three changes must have occurred in separate stages, following Kawakami (1995, as cited by Nakai 2001 and Uwano 2012) who posits that the gain of initial H occurred before the change of final H/F to L. I posit that the third change occurred later in a separate stage, following Nakai (2001)'s timeline of the development of Modern Kyoto and Frellesvig (2010), who writes that the change from HHL to HLL was completed recently.

Conversely, the correspondences in (36) show that Classes 1, 6, and 7 did not develop into initial accent. Classes 1 and 7 remain the same from EMJ to Modern Kyoto, while Class 6, like Class 2 above, loses its multiply linked H tones, a change which I posit occurred prior to the change of Class 2 HHL to HLL (change 3 in the previous paragraph), following Nakai (2001) who places LHH > LLH before HHL > HLL in his development timeline. Note that these classes share neither of the characteristics described above for either Class 2 and 3, which begin with H and have an HL pitch fall, or Classes 4, 5a, and 5b, which begin with two L toned moras. That is, while Class 1 begins with H, it does not have an HL pitch fall, and neither Class 6 nor Class 7 begin with two L toned moras.

Accordingly, I posit four stages of development from EMJ to Modern Kyoto, summarized below. Stage 1 involves gain of initial H where applicable (Classes 4, 5a, and 5b), stage 2 involves deletion of a second prominence (enforcement of culminativity, a restriction against multiple separate H tones in a single word) where applicable (Classes 5a and 5b), stage 3 involves peak delay (Class 6), and stage 4 involves a restriction against multilinked H tones (Classes 2 and 4). These developments are summarized in the table in (37) below. Cells shaded in gray indicate that a form has changed from the previous stage, indicated in the cell to the immediate left of the shaded cell.

(37) Development from EMJ to Modern Kyoto

Class	EMJ	Initial H Gain	Culminativity	Peak Delay	No Multiply Linked Hs	Modern Kyoto
1	HHH	HHH	HHH	HHH	HHH	HHH
2	HHL	HHL	HHL	HHL	HLL	HLL
3	HLL	HLL	HLL	HLL	HLL	HLL
4	LLL	HLL ~ HHL	HLL ~ HHL	HLL ~ HHL	HLL (~ HHL)	HLL (~ HHL)
5a	LLH	HLH	HLL	HLL	HLL	HLL
5b	LLF	HLF	HLL	.HLL	HLL	HLL
6	LHH	LHH	LHH	LLH	LLH	LLH
7	LHL	LHL	LHL	LHL	LHL	LHL

Note that although the first stage in this development is Initial H Gain wherever applicable (that is, whenever a word begins with two L tones), once Culminativity comes into play in Stage 2, remaining words beginning with two L tones (belonging to Class 6 and starting in the Peak Delay column, as shown in (37) above), will *not* gain an initial H again, as such a gain would output HLH in violation of culminativity, and a change to HLL would violate faithfulness constraints. The fact that constraint promotions earlier in the developmental pathway can have consequences on possible output forms in future stages is discussed and demonstrated in Section 5.

That the development must have proceeded in multiple stages is confirmed by additional documents recording the changing accent between the *Ruijumyogisho* and Modern Kyoto. These documents are the *Shizakoshiki*, written in the Kamakura era (12th – 14th centuries), a musical score of Buddhist vocal music which is said to reflect the Kyoto Japanese accentual system of the time (as Japanese traditional vocal music is said to often reflect accent), and the *Bumoki*, written in the early Edo period (17th-19th centuries), a reference book of Buddhist chants which also recorded the Kyoto Japanese accentual system of the time (Nakai 2001). The accentual patterns of three mora nouns from EMJ to Modern Kyoto including these two stages, along with a separation between Early Edo and Late Edo, to each of which Nakai (2001) attributes different changes, are given in the table in (38) below. While there appear to be no changes between EMJ and the Kamakura era, there were changes between the Kamakura and Edo eras, the Early Edo era and the Late Edo era, and the Edo era and the present time. Cells shaded in gray indicate that a form has changed from the previous stage, indicated in the cell to the immediate left of the shaded cell. Classes 2 through 5b, enclosed in bolded borders below, will develop into HLL by Modern Kyoto.

(38) Accentual patterns of three mora nouns from EMJ to Modern Kyoto (Nakai 2001)

Class	EMJ (*Ruijumyogisho*)	Kamakura (*Shizakoshiki*)	Early Edo (*Bumoki*)	Late Edo	Modern Kyoto
1	HHH	HHH	HHH	HHH	HHH
2	HHL	HHL	HHL	HHL	HLL
3	HLL	HLL	HLL	HLL	HLL
4	LLL	LLL	HHL	HHL	HLL ~ HHL
5a	LLH	LLH	HLL	HLL	HLL
5b	LLF	LLF	HLL	HLL	HLL
6	LHH	LHH	LHH	LLH	LLH
7	LHL	LHL	LHL	LHL	LHL

The developmental pathway proposed here focuses on developmental trends as suggested by the correspondences listed in (35) and (36) above, documents such as the *Shizakoshiki* and the *Bumoki*, and databases comparing EMJ accent to Modern Kyoto accent such as Sugito (1995)'s dictionary. However, it should be noted that not all words follow the developmental trends summarized above, developing into other forms instead. Table (35) shows that Class 4 (LLL) words may also develop into HHL instead of HLL, though this distinction was eventually leveled, causing a merger to HLL in recent Modern Kyoto, per Frellesvig (2010). In addition, although the trend was for Classes 2 (HHL), 5a (LLH), and 5b (LLF) to become HLL as shown in (35), a small amount of words from these classes became LHL (notably a non-initial accent pattern), HHH, and HHL respectively instead, as shown in their respective rows in (36). Thus, developmental correspondences cannot always be viewed as "EMJ Form X will always become Modern Kyoto Form Y." Instead, correspondences should be understood to be developmental trends. These divergent developments are discussed below, alongside the developmental trends discussed above.

In the next section, I discuss each of the four developmental stages in turn, beginning from EMJ and proceeding to Modern Kyoto. Stage 1, Initial H Gain, is discussed in Section 5.1; Stage 2, Culminativity, is discussed in Section 5.2; Stage 3, Peak Delay, is discussed in Section 5.3; and Stage 4, No Multiply Linked Hs, is discussed in Section 5.4. The divergent developments described in the previous paragraph will also be discussed – the divergent development in Class 4 (LLL → HLL or HHL) is discussed with Stage 1 in Section 5.1, the divergent developments in Classes 5a (LLH → HLL or HHH) and 5b (LLF → HLL or HHL) are discussed with Stage 2 in Section 5.2, and the divergent development in Class 2 (HHL → HLL or LHL) is discussed with Stage 4 in Section 5.4.

5 The Road to the Modern Kyoto Accentual System

The proposal discussed in this section accounts for the changing grammar of the Kyoto branch accentual system in Optimality Theory (Prince and Smolensky 1993/2004). Framing this development in Optimality Theory allows for the precise specification of the state of the grammar at each stage and of the ways in which each stage differs through the reranking of both universal constraints and a language-specific constraint (INITIAL-H). This framing also allows us to see how changes in the grammar in earlier stages can have consequences for developments in later stages.

The account presented below uses the following constraints in (39) and a few additional constraints to be discussed below.

(39) List of constraints in the present analysis

a. **INITIAL-H (INIT-H)**: A prosodic word must begin with an H tone.
Assign one violation for every prosodic word which does not have an H tone associated with its first mora.

b. **NOMULTILINK-H (NML-H)**: An H tone must not be associated to more than one mora.
Assign one violation for every additional mora above 1 to which H is linked.

c. **IDENT-IO(TONE) (ID-TONE)**: Moras and tones must have the same correspondence relationships in the output as they do in the input.
Assign one violation for every mora-tone pair which correspond in the input that do not correspond in the output.

d. **MAX-TONE-H (MAX-H)**: Do not delete H tones.
Assign one violation for every H tone in the input which is not present in the output.

e. **MAX-TONE-L (MAX-L)**: Do not delete L tones.
Assign one violation for every L tone in the input which is not present in the output.

f. **OBLIGATORY CONTOUR PRINCIPLE (OCP)**: Adjacent identical tones are prohibited.

Assign one violation for every pair of adjacent identical, but separate tones.

g. **CULM(INATIVITY)**: A prosodic word must not have more than peak.

Assign one violation for every prosodic word which has more than one H tone.

h. **NOCONTOUR-μ (NC-μ)**: A mora should not have a tonal contour.

Assign one violation for every mora which is associated to more than one tone.

i. **DEP-TONE**: Do not insert a tone.

Assign one violation for every prosodic word which has a tone in the output which it did not have in the input.

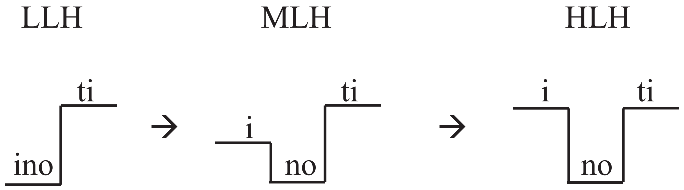
5.1 Stage 1: Initial-H

Naturally, it seems unlikely that speakers suddenly posited that more words should have an initial H tone without any motivation. Thus, a logical question arises: Why should initial accent gain occur? I propose that the answer to this question lies in an increase in the frequency of initial H tones due to a pitch rise enhancement process.

5.1.1 Enhancement of a Pitch Rise Led to Increased Surface Frequency of Initial H

Kawakami (1995, as cited by Nakai 2001 and Uwano 2012) proposed that in the course of the development from EMJ to Modern Kyoto, L tones immediately preceding H tones were lowered in order to make the pitch rise from L to H more acoustically salient by enhancing the contrast between L and H. Kawakami proposed that when this occurred, if the mora with the lowered L tone was preceded by other L-toned moras, the lowering would cause all preceding L tones to drift upward in pitch to something like a mid tone (M). Eventually, these mid tones drift even higher until they are H tones. In addition, learners can interpret the presence of a fall from M to L as the presence of an accent, further propelling the upward drift to H over the course of Modern Kyoto's development. This is schematized in (40) below with the word *inoti* 'life,' which in EMJ belongs to class 3.5a and had the accentual pattern LLH. Lines indicate pitch levels, with the lowest level corresponding to L, the middle level corresponding to M, and the highest level corresponding to H.

(40) Pitch rise enhancement in an LLH word inoti 'life'; LLH → MLH → HLH



By this contrast enhancement process, forms like LLH and LLF receive an initial H. However, the frequencies of LLH and LLF are low, at 11.1% and 1.7%, respectively, so this cannot be the full story. Recall from the table of correspondences in (35) above that LLL eventually develops into HLL as well. Unlike LLH and LLF, however, LLL does not have an H tone within it that could prompt such a pitch rise enhancement to occur. Since an LH sequence is necessary for the process to occur, where does the requisite H tone come from? The answer to this lies in common particles which can follow words, such as the topic marker *wa*.

In modern Japanese, common particles never have their own independent tone, taking on the same tonal value as the mora immediately preceding it, unless the word has word-final accent. Accordingly, if a word ends in an H tone, a following particle will also have an H tone, e.g. *SAKURA-GA* 'cherry (Nom.)' Similarly, if a word ends in an L tone, a following particle will also have an L tone, e.g. *kaBU'to-ga* 'helmet (Nom.).' When words have a final accent, the particle instead takes on the L tone of the accentual HL complex, e.g. *osaKI'-ga* 'future (Nom.).' Although the particle does not take on the same tonal value as the mora immediately preceding it in this case, the tone nonetheless is received from the accentual complex associated with the final mora of the lexical word and is not underlyingly associated to the particle.

However, in EMJ, particles are independent accentual units with their own underlying H tone, using the same accentual system as nouns (Nakai 2001, Frellesvig 2010). Therefore, common monomoraic particles such as the genitive marker *ga*[12], the accusative marker *wo*, and the topic marker *wa* have the accentual pattern H, and a common bimoraic particle such as the ablative marker *kara* has the accentual

---

[12] The modern Japanese nominative marker *ga* descends from the Old Japanese genitive marker *ga*. This change occurred in the Late Middle Japanese period (13th – 16th centuries) (Frellesvig 2010).

pattern HH (Frellesvig 2010). Thus, when these H-toned particles are attached to a word with an accentual pattern with no H tones, such as LL and LLL, the pitch rise enhancement process described above can occur as expected, because an H tone now exists which can prompt the process to occur. Thus, an LL word like *ike* ‘pond’ will first pass through ML by upward drift of the first L, and eventually become HL by further upward drift. Likewise, an LLL word like *otoko* ‘man’ will pass through MML and eventually become HHL by further upward drift, as all L tones preceding the lowered L tone will drift upward. As such, in addition to LLH and LLF which become HLH and HLF due to this contrast enhancement process, LL-H becomes HL-H, and LLL-H becomes HHL-H, where the tone after the hyphen represents an H-toned particle.

The application of the pitch rise enhancement process to these forms increases the prevalence of initial H tones. In bimoraic words, the original prevalence of initial H tones was 42.4% (30.4% HH + 11.6% HL + 0.4% HF), as shown in (27) and (28). The change of LL to HL increases that prevalence to 76.2% (42.4% + 33.8% LL) in bimoraic words. In trimoraic words, the original prevalence of H tones was 43.8% (31.7% HHH + 8.3% HHL + 2.1% HLL), as shown in (24) and (25). The change of LLH to HLH, LLF to HLF, and LLL to HHL (by the pitch enhancement process; in the following paragraphs, I argue that LLL could also become HLL in this stage) increases that prevalence to 86.7% (43.8% + 11.2% LLH + 1.7% LLF + 30.0% LLL). Given the increased prevalence of H, I propose that new learners phonologize this increased prevalence by positing a constraint requiring that words have an initial H. This constraint is given in (41) below.

(41) INITIAL-H: A prosodic word must begin with an H tone.

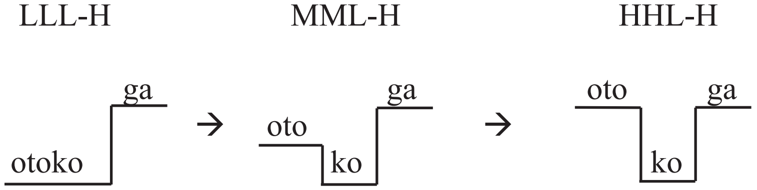
Assign one violation for every prosodic word which does not have an H tone associated with its first mora.

I return to an Optimality Theoretic account using INITIAL-H in Section 5.1.2. For now, let us consider the application of pitch rise enhancement to the EMJ form LLL, which requires more comment, as LLL has three L-toned moras. First, recall from (35) above that the LLL accentual pattern in EMJ corresponds to HLL ~ HHL in MKJ. I argue that LLL could change to either HLL or HHL in this stage. HLL results from application of INITIAL-H to LLL, while HHL results from application of the pitch rise enhancement process to LLL. First, let us consider the latter pitch rise enhancement pathway.

If the pitch rise enhancement process applies as described above, then HHL results transparently, as the upward drift of L tones to H will apply to all L tones preceding the lowered L which immediately precedes the inherent H tone of the particle. This is shown in (42) below.

(42) Pitch rise enhancement in an LLL word with H-toned particle

otoko-ga ‘man (Nom.)’; LLL-H → MML-H → HHL-H



Examination of words recorded in the *Bumoki* from the Nanbokucho era (14th century) reveals correspondences between LLL and HHL, so it is known that a change from LLL to HHL must have taken place and that Kawakami’s proposed pitch rise enhancement process can naturally account for the change (Nakai 2001). Whereas initial H insertion in LLH, LLF, and LL may cause a learner to posit a constraint requiring initial H, a learner could posit that they heard a high beginning and second-syllable accent in a word like *otoko* and posit the underlying form HHL (and also obtaining more proof for initial H).

On the other hand, how does HLL result from LLL? I propose that it can emerge in one of two ways. The first way is that LLL became HHL by this pitch rise enhancement process. A learner posits a high beginning and second mora accent as the word’s underlying form, and in a later stage (Stage 4 in the present proposal), a restriction against multilinked H deletes association from the H tone to the second mora, leaving HLL. Such a pathway is confirmed as a possibility: as noted by Nakai (2001), Frellesvig (2010) and examination of several words in Sugito (1995), HHL patterns (which were originally LLL in EMJ) in the older generations are increasingly becoming HLL in the younger generations. An example is *otoko* ‘man,’ which was

LLL in EMJ, HHL for all three older generation speakers recorded by Sugito (1995) who were all born between 1916 and 1932, and HLL for all three younger generation speakers recorded by Sugito who were all born between 1962 and 1964.

The restriction against multilinked H may also create LHL from HHL originating from LLL. An example of a word showing this possibility is *kagami* ‘mirror,’ which is pronounced as LHL by one of Sugito (1995)’s six speakers. Older generation speakers 1 and 3 (born 1932 and 1916 respectively) produce *kagami* as HHL, while older generation speaker 2 (born 1923) produces it as HLL. Younger generation speakers 1 and 2 (born 1964 and 1962 respectively) produce it as HLL, while younger generation speaker 3 (born 1963) produces it as LHL. *Azuma* ‘east,’ another word which was LLL in EMJ, was produced as LHL by all six speakers in Sugito (1995).

A second possible way for LLL to have developed into HLL is that while some learners posited that the pitch rise-enhanced HHL form they heard had second syllable accent, other learners instead applied the INITIAL-H constraint to LLL forms they heard and posited HLL as the underlying form instead. Examination of other originally LLL words in Sugito (1995) reveals that it is not always the case that the older generation has HHL and the younger generation has HLL. Sometimes, the older generation has HLL, while the younger generation has HHL, as is the case for *hakama* (an article of traditional Japanese clothing resembling pants). Other times, both the older generation and the younger generation have HLL, as is the case for *hikari* ‘light.’ While about 300 years have passed between the dialect recorded in the *Bumoki* and the recording of the data in Sugito’s dictionary (thus making it possible that even these words had originally become HHL from LLL in an earlier, though unrecorded, period), it is also possible that learners had simply posited that the underlying form was HLL due to INITIAL-H. For the present analysis, I posit that both developmental pathways happen.

Whether LLL becomes HHL as a result of the pitch rise enhancement process or HLL by application of INITIAL-H, the end result is the same with respect to initial H frequencies: the surface prevalence of word-initial H increases to 86.7% in trimoraic words as discussed above. Considering bimoraic and trimoraic nouns together, a significant proportion of nouns – 80.8% of bimoraic and trimoraic nouns combined – appear to have initial H to learners. Even if LLL only becomes HHL in a first stage as a result of the pitch rise enhancement process, this still gives learners a large set of inputs from which to posit an INITIAL-H constraint and then to later apply that constraint to LLL words to obtain HLL.

If INITIAL-H is active, then why do words with the forms LH, LF, LHH, and LHL resist initial H gain? I argue that this is due to two constraints: 1) MAX-TONE-L, a constraint against deleting L tones and 2) CULMINATIVITY, a constraint against a word having two adjacent identical tones – in this case, two adjacent identical H tones. The effects of these two constraints is discussed in Section 5.1.2.

The pitch rise enhancement process may also have contributed to greater incidence of peak delay, which I posit will eventually cause LHH to change to LLH. This is due to the fact that the enhancement process lowers an L tone preceding an H tone, increasing the pitch distance between the two targets and, due to articulatory constraints on pitch changes which make pitch rises require more effort than pitch falls (Xu and Sun 2002), increasing the chances that peak delay will occur when pitch rise enhancement has also occurred. More will be said about this in section 5.3.

5.1.2 Initial H Gain in Optimality Theory

As discussed above, a pitch rise enhancement process applied to EMJ forms beginning with two or more adjacent L toned moras followed by an H toned mora, increasing the surface prevalence of initial H to 76.2% in bimoraic words and 86.7% in trimoraic words. This increase leads learners in a new generation to posit the INITIAL-H constraint requiring that words begin with an H toned mora.

Because this will change the identity of the tone linked to the first mora and insert a new H tone, I propose that it must be ranked above IDENT-IO(TONE) and DEP-TONE. Their ranking is given in (43) below.

(43) INITIAL-H >> IDENT-IO(TONE), DEP-TONE

To illustrate how IDENT-IO(TONE) works, consider the following examples. For this paper, standard Autosegmental Theoretic (Goldsmith 1976) representations will be assumed such that moras bearing adjacent tones are assumed to be linked to a single tone of that value. Correspondences between moras and tones are indicated by subscripts. Matching subscripts indicate correspondence. Because of the assumed autosegmental representation, tones may have more than one correspondence. The example in (44) incurs no violations of IDENT-IO(TONE), because correspondences between moras and tones in the input are retained in the output.

(44) Input: LLH Output: LLH

The example in (45), however, incurs one violation of IDENT-IO(TONE) because the correspondence between the first mora and the L tone is not maintained, and mora 1 (μ_1) now corresponds with an H tone. This change also constitutes a DEP-TONE violation, as a new tone which did not exist in the input has been inserted in the output.

(45) Input: LLH Output: HLH

As shown in the examples above, correspondence subscripts reflect autosegmental association lines. As such, I will use correspondence subscripts to indicate association lines in the tableaux below.

The EMJ forms of Classes 4 LLL, 5a LLH, and 5b LLF will gain an initial H because of INITIAL-H. Classes 6 LHH and 7 LHL will not gain an initial H because an output respecting INITIAL-H (as in (46) below) would incur violations of MAX-TONE-L, which prohibits the deletion of an L tone.

(46) Input: LHL Output: HHL

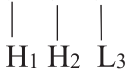
A similar loss of H would incur a violation of MAX-TONE-H, which prohibits the deletion of an H tone, preventing Classes 6 and 7 from changing to HLL. This is shown in (47) below, in which both L_1 and H_2 are lost in the output.

(47) Input: LHL Output: HLL

These two constraints generally do not interact with each other, and their relative ranking in Stages 1, 3, and 4 will not select between candidates. However, I propose in Stage 2 that the constraints are freely rankable, with consequences for the output depending on their ranking in the grammar selected by a learner.

In addition, the initial L tone of Classes 6 LHH and 7 LHL will not simply change to an H tone in order to satisfy INITIAL-H. Such a change in tonal identity will incur a violation of the OBLIGATORY CONTOUR PRINCIPLE (OCP), which states that adjacent identical tones are prohibited. Because these forms already have an H tone associated with their second mora, change of the initial L to H will cause two identical, but separate, tones, here H_1 and H_2, to come into contact with each other, as shown in (48) below.

(48) Violation of the OCP due to initial H insertion into LHL



The MAX-TONE constraints and OCP must be ranked above INITIAL-H, as will be shown in the tableaux in (49) through (53). They are unranked with respect to each other.

The EMJ forms of Classes 1 HHH, 2 HHL, and 3 HLL will also not gain an initial H because their underlying forms already have an initial H and thus do not violate the INITIAL-H constraint.

The tableaux in (49) through (53) below account for initial H gain in Stage 1. As stated above, LLL, LLH, and LLF will gain an initial H. This is illustrated in the tableau in (49) through (51). The falling tone is represented as [HL] in the tableau below in order to make reference to its internal structure. In addition, for this and all following tableaux, I assume an unviolated LINEARITY-TONE constraint, which requires the ordering of tones in the output to retain their ordering in the input. Accordingly, any candidates which appear to have modified the input ordering I assume have deleted an input tone(s) and inserted new tone(s) to yield the new ordering. Thus, I assume for example that (49c) below has deleted the original tone labeled H_3 and inserted a new tone labeled H_1.

(49) LLH > HLH[13]

/$L_{1,2}H_3$/ *àbùrá* 'oil'	MAX-H	MAX-L	OCP	INIT-H	NC-μ	ID-TONE	CULM	DEP-TONE
☞ a. $H_1L_2H_3$ *ábùrá*						* (μ_1)	*	* (H_1)
b. $L_{1,2}H_3$ *àbùrá*				*! W		L	L	L
c. $H_1L_{2,3}$ *ábùrà*	*! W (H_3)					** W (μ_1,μ_3)	L	* (H_1)

---

[13] Following standard usage in historical linguistics, the > symbol used here signifies that LLH "changed into" HLH as a result of this development (Campbell 2001).

(50) LLL > HLL

/$L_{1,2,3}$/ *òtòkò* ‘man’	Max-H	Max-L	OCP	Init-H	NC-μ	Id-Tone	Culm	Dep-Tone
☞ a. $H_1L_{2,3}$ *ótòkò*						* ($μ_1$)		* (H_1)
b. $L_{1,2,3}$ *òtòkò*				*! W		L		L

(51) LL[HL] > HL[HL]

$/L_{1,2}[HL]_3/$ *tàmàkî* ‘elbow guard’	Max-H	Max-L	OCP	Init-H	NC-μ	Id-Tone	Culm	Dep-Tone
☞ a. $H_1L_2[HL]_3$ *támàkî*					*	* (μ_1)	*	* (H_1)
b. $L_{1,2}[HL]_3$ *tàmàkî*				*! W	*	L	L	L
c. $H_1L_{2,3}$ *támàkì*	*! W (H_3)				L	** W (μ_1,μ_3)	L	* (H_1)
d. $H_1L_2L_3$ *támàkì*	*! W (H_3)		*! W		L	** W (μ_1,μ_3)	L	* (H_1)
e. $H_1L_2H_3$ *támàkí*		*! W (L_3)			L	** W (μ_1,μ_3)	*	* (H_1)

As shown in the tableaux, the faithful candidates in each case are ruled out by INITIAL-H. MAX-TONE-H rules out changing to initial accented HLL in one stage (49c, 51c), with their single violations of MAX-TONE-H being incurred by the deletion of the H tone which was originally associated with the third mora. Although GEN can generate candidates which are any number of changes away from the input (in this case, two changes: the change of the first mora to H and the change of the last mora to L), it must be kept in mind that the changes in the development of Modern Kyoto are not synchronic but diachronic and occurred over several stages. Therefore, the initial accented candidates in (49c, 51c) must crucially be ruled out at this stage in the development of Modern Kyoto because only initial H is gained at this stage. The second step, deletion of a second prominence, occurs in a later stage of this development.

Additionally, the resulting HLH and HLF forms crucially survive this stage because CULMINATIVITY is ranked below INITIAL-H, and MAX-TONE-H rules out LLH > HLL and LLF > HLL at this stage. The ranking of INITIAL-H above DEP-TONE at this stage also allows for the insertion of an initial H. As the tableau shows, IDENT-IO(TONE), CULMINATIVITY, and DEP-TONE must be ranked below INITIAL-H, though a ranking cannot be established between them. Furthermore, while the exact position of NOCONTOUR-μ cannot be determined from the contests shown in above, it must be ranked relatively low, at least below MAX-TONE-H and MAX-TONE-L, as the candidate preserving the falling contour in (51a) wins over the candidates which simplify the contour in (51c-e), which all fare worse than the winning candidate on a combination of MAX-TONE-H, MAX-TONE-L, and the OBLIGATORY CONTOUR PRINCIPLE. The tableau above thus ranks NOCONTOUR-μ as high as possible, in the same stratum as and unrankable with respect to INITIAL-H. NOMULTILINK-H, which is not shown, plays no role in (49) through (51), as no moras are multiply-linked with an H tone.

LHH and LHL will *not* gain an initial H. This is illustrated in the tableaux in (52) and (53).

(52) LHH > LHH

/$L_1H_{2,3}$/ *ùságí* ‘rabbit’	Max-H	Max-L	OCP	Init-H	NC-μ	Id-Tone	Culm	Dep-Tone	NML-H
☞ a. $L_1H_{2,3}$ *ùságí*				*					*
b. $H_1H_{2,3}$ (L_1 deleted, H_1 inserted) *úságí*		*! W (L_1)	*! W	L		* W ($μ_1$)	* W	* W (H_1)	*
c. $H_1H_{2,3}$ (L_1 changed to H_1) *úságí*			*! W	L		* W ($μ_1$)	* W		*
d. $H_1L_{2,3}$ *úsàgì*	*! W ($H_{2,3}$)	*! W (L_1)		L		*** W ($μ_1$, $μ_2$,$μ_3$)		** W (H_1, $L_{2,3}$)	L
e. $H_{1,2,3}$ *úságí*		*! W (L_1)		L		* W ($μ_1$)			** W
f. $L_{1,2}H_3$ *ùsàgí*				*		* W ($μ_2$)			L

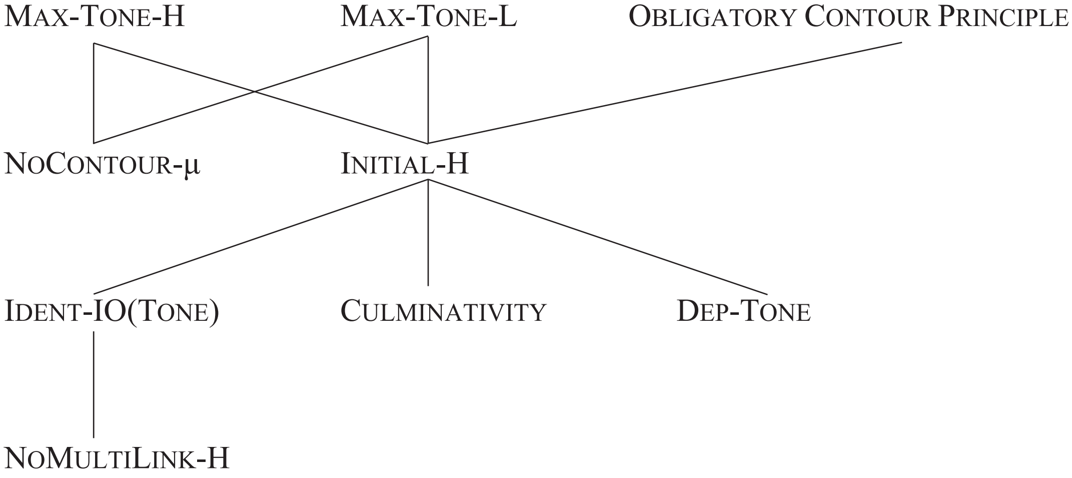
(53) LHL > LHL

/$L_1H_2L_3$/ *kàbútò* ‘helmet’	Max-H	Max-L	OCP	Init-H	NC-μ	Id-Tone	Culm	Dep-Tone	NML-H
☞ a. $L_1H_2L_3$ *kàbútò*				*					
b. $H_1H_2L_3$ (L_1 deleted, and H_1 inserted) *kábútò*		*! W (L_1)	*! W	L		* W (μ_1)	* W	* W (H_1)	
c. $H_1H_2L_3$ (L_1 changed to H_1) *kábútò*			*! W	L		* W (μ_1)	* W		
d. $H_1L_{2,3}$ *kábùtò*	*! W (H_2)	*! W (L_1)		L		** W (μ_1,μ_2)		* W (H_1)	
e. $H_{1,2,3}$ *kábútó*		*!* W (L_1, L_3)		L		** W (μ_1,μ_3)			** W
f. $H_{1,2}L_3$ *kábútò*		*! W (L_1)		L		* W (μ_1)			* W

Although the faithful candidates LHH (52a) and LHL (53a) do not have initial Hs, they are nonetheless selected as the winners. (52b, e) and (53b, f) violate MAX-TONE-L due to deleting the original L tone associated with the first mora. (52d, 53d) additionally delete an H tone, resulting in one violation each of MAX-TONE-H. (53e) violates MAX-TONE-L twice due to deleting both L tones present in the input. In the case of (52d), both original tones are deleted and replaced by new tones (incurring violations of the low-ranked DEP-TONE). (52d, e) and (53d, f) demonstrate that both MAX-TONE-H and MAX-TONE-L are crucially ranked above INITIAL-H. The candidates (52c) and (53c) differ from the candidates in (52b) and (53b) in that the latter pair obtained their H_1 tones by deletion of L_1 and insertion of a new H tone, while the former pair obtained their H_1 simply by changing the tonal identity of L_1. (52c) and (53c), which do not violate either of the MAX-TONE constraints, are instead ruled out by the OBLIGATORY CONTOUR PRINCIPLE, which, as shown above, crucially dominates INITIAL-H. Finally, candidate (52f), which only has one association line with an H tone, is ruled out by IDENT-IO(TONE), which crucially dominates NOMULTILINK-H.

The tableaux above demonstrate the constraint hierarchy represented in the Hasse diagram in (54) below.

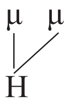
(54) Stage 1 Constraint Hierarchy



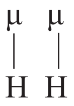
5.1.3 How Does LL become HL under standard Autosegmental Theoretic Assumptions?

As noted above, in the present analysis, I assume the standard Autosegmental Theoretic representation for tonal association. Specifically, I assume that adjacent moras bearing an identical tone value are linked to a single tone, as shown in (55) below, not to two separate tones, as shown in (56) below.

(55) Adjacent identical tones are linked to a single tone…



(56) …not to two separate tones.



Accordingly, an LL word like *ike* ‘pond’ has the following associations in (57).

(57) Tonal associations for *ike* ‘pond’

ike
|/
L

If this is the case, then how can a pitch rise enhancement process result in the second L lowering when immediately followed by an H tone (such as that supplied by a common particle like the genitive *ga*) and the first L raising, eventually to H, creating a separate tone? Based on the representation, since there is only one L tone linked to two moras, we might expect the value of both moras to be lowered in the contrast enhancement process.

This process can be understood as a process that concerns primarily articulation and articulatory constraints in a system which is necessarily implemented and linearly ordered in time. In a phrase like *ike-ga* ‘pond (Nom.)’, which has the tonal pattern LL-H, there are two pitch targets, low and high. At the beginning of the utterance, musculature involved in pitch lowering (the sternohyoid muscle, according to Sugito (2001)) is activated, resulting in a low pitch on the first mora. In order to actuate a pitch rise, musculature involved in pitch raising (the cricothyroid, according to Sugito (2001)) are activated. In order for the high pitch target to be reached in the correct syllable, activation of this musculature must start earlier than the syllable in question. This is because of a physiological constraint which makes executing pitch rises take more time than executing pitch falls (Xu and Sun 2002). This means that activation of the cricothyroid must start during the preceding syllable, the second L toned syllable. However, since that syllable is already associated with an L tone target, activating the cricothyroid during the second syllable will cause its pitch to begin raising in order to reach the H tone target in the next syllable. In order to prevent this second syllable from drifting away from the L toned target, the speaker implements greater activation of the sternohyoid in order to keep pitch low and counter the raising action of the cricothyroid. Finally, during the H toned syllable, only the cricothyroid, responsible for pitch raising, remains activated. The diagram in (58) displays the muscles which are active during each part of the utterance.

(58) Muscles active in the production of *ike-ga* ‘pond (Nom.)’

Syllable	i	ke	ga
Tone	L	L	H
Mechanisms Involved	Pitch Lowering	Preparation for pitch rise; enhanced pitch lowering	Pitch raising
Muscles Involved	Sternohyoid	Cricothyroid; Sternohyoid	Cricothyroid

Because these mechanisms are necessarily implemented in time, and because pitch raising is not yet needed in the first syllable, the activation of the cricothyroid will not begin until the second syllable. This means that the sternohyoid activation does not need to be increased until the second syllable. This results in uneven sternohyoid activation between the first and the second syllables, with greater activation in the second syllable than in the first syllable, resulting in the pitch being lower in the second syllable than in the first syllable, despite both syllables being linked to the same L tone, which will result in upward pitch drift of the first syllable, which is then perceived by learners as being an initial H tone.

Summarizing this section, I propose that the pitch rise enhancement process increased the surface prevalence of initial H, which led new generations of learners to posit an INITIAL-H constraint, causing initial accent gain. The next section discusses culminativity.

5.2 Stage 2: Culminativity

5.2.1 Relatively Low HLH Frequencies Contribute to Culminativity

At the end of stage 1, initial H gain, the following changes have taken place, shown in (59).

(59) Stage 1 (Initial H Gain) Changes

Class	EMJ	End of Stage 1
4	LLL	HLL ~ HHL
5a	LLH	HLH
5b	LLF	HLF

By the time of the writing of the *Bumoki* in the late 17th century, Class 4 had remained unchanged since the end of stage 1, and Classes 5a and 5b had become HLL, losing the second high tone on their third mora. This is due to a restriction against the presence of multiple peaks in a word. Called “single peaking” in Kindaichi (1973, cited by Nakai 2001) and Matsumori (1999) and PLATEAU by Cassimjee and Kisseberth (1998, cited by Kisseberth 2001), I will refer to this restriction here as “culminativity,” following Kubozono (2012)’s usage of this term for Japanese. I use the same term to refer to the constraint involved in this stage, CULMINATIVITY, which I posit to have been promoted to a higher position in the constraint hierarchy due to the relatively low frequency of forms with two separate H tones, as will be seen below.

CULMINATIVITY can be understood to be a type of OCP effect restricting the number of top-level prominences (here, H tones) in a word to a single instance, as schematized in a metrical grid as introduced by Liberman (1975) for English stress and further developed by Liberman and Prince (1977), Prince (1983), and Halle and Vergnaud (1987). This is shown in (60) and (61) below with two grid tiers. The example in (60) violates CULMINATIVITY because two grid marks, represented as asterisks (*) following Halle and Vergnaud (1987)'s notation, occupy the highest tier in the metrical grid, while the examples in (61) satisfy CULMINATIVITY, as only one grid mark occupies the highest tier in the grid.

(60) Metrical Grid Violating CULMINATIVITY

```
*       *
*   *   *
H   L   H
```

(61) Metrical Grids Satisfying CULMINATIVITY

```
*                 *                 *
*   *   *     *   *   *     *   *   *
H   L   L     L   H   L     L   L   H
```

While the forms HLH and HLF are never attested in three mora nouns in any historical record of accent in Japanese,[14] Nakai (2001) states that words with multiple

---

[14] Interestingly, younger generation speaker 1 in Sugito (1995) produces the word *tasuki* 'cord used to tuck up the sleeves of a kimono' as HLH. *Tasuki* was LLH in EMJ. It is the only example of HLH in a noun produced by a Modern Kyoto speaker in Sugito (1995). The other speakers produce *tasuki* as HLL (2 speakers), HHL (1 speaker), LHL (1 speaker), and HHH (1 speaker).

Also of note is that all three older generation speakers produce the adverb *shikamo* 'moreover,' which was LLF in EMJ, as HLH, and older generation speaker 2 produces the adverbial *ika ni* 'no peaks separated by an L tone would have existed in late Heian era EMJ (11th to 12th centuries) due to particles bearing their own independent lexical H tone rather than being tonally subordinate to preceding material, as particles are in Modern Japanese. Examples are given in (62) below with the high-pitched topic particle *wa*. Bimoraic xamples (59a-d) are from Nakai (2001). Trimoraic examples (e-i) were constructed from three mora words taken from Sugito (1995) and the topic particle *wa*.

(62) HL_0H in late Heian era EMJ

a. *isi* ‘stone’ → *isi-wa* ‘stone (topic)’

HL → HL-H

b. *ame* ‘rain’ → *ame-wa* ‘rain (topic)’

LF → LF-H

c. *mizo* ‘ditch’ → *mizo-wa* ‘ditch (topic)’

HF → HF-H

d. *hagi* ‘leg’ → *hagi-wa* ‘leg (topic)’

RL → RL-H

e. *tobira* ‘door’ → *tobira-wa* ‘door (topic)’

HHL → HHL-H

f. *hatake* ‘cultivated land’ → *hatake-wa* ‘cultivated land (topic)’

LHL → LHL-H

---

matter how…’ as HLH as well, matching its EMJ form of HLH. All three younger generation speakers produce *shikamo* and *ika ni* as LHL.

g. *tikara* ‘power’ → *tikara-wa* ‘power (topic)’

HLL → HLL-H

h. *yorozu* ‘ten thousand’ → *yorozu-wa* ‘ten thousand (topic)’

LLF → LLF-H

i. *zisyaku* ‘magnet’ → *zisyaku-ha* ‘magnet (topic)’

RLL → RLL-H

Nouns like (62a-d) account for 18.9% (= 11.6% HL + 6.5% LF + 0.4% HF + 0.4% RL) of all two mora nouns in Sugito (1995), while nouns like (62e-i) account for 18.6% (= 8.3% HHL + 6.0% LHL + 2.1% HLL + 1.7% LLF + 0.5% RLL) of all three mora nouns in Sugito (1995), yielding a prevalence of HL_0H in late Heian era EMJ two to three mora nouns of 18.8%.

According to Nakai, particles became subordinate to preceding material between the Heian era and the Kamakura era. I posit that this is due to peak delay, discussed below. Kamakura era accentual patterns are given in (63), corresponding to the late Heian era accentual patterns in (62). Additionally, word-initial R tones change to H in this period (Nakai 2001), so *hagi* changes from RL to HL, and *zisyaku* changes from RLL to HLL.

(63) Particles tonally subordinate to preceding material in Kamakura era EMJ

a. *isi* ‘stone’ → *isi-wa* ‘stone (topic)’

HL → HL-L

b. *ame* ‘rain’ → *ame-wa* ‘rain (topic)’

LF → LF-L

c. *mizo* ‘ditch’ → *mizo-wa* ‘ditch (topic)’

HF → HF-L

d. *hagi* ‘leg’ → *hagi-wa* ‘leg (topic)’

HL → HL-L

e. *tobira* ‘door’ → *tobira-wa* ‘door (topic)’

HHL → HHL-L

f. *hatake* ‘cultivated land’ → *hatake-wa* ‘cultivated land (topic)’

LHL → LHL-L

g. *tikara* ‘power’ → *tikara-wa* ‘power (topic)’

HLL → HLL-L

h. *yorozu* ‘ten thousand’ → *yorozu-wa* ‘ten thousand (topic)’

LLF → LLF-L

i. *zisyaku* ‘magnet’ → *zisyaku-wa* ‘magnet (topic)’

HLL → RLL-L

The tonal subordination of particles may have occurred due to peak delay, a phenomenon in which F0 peaks occur later than the syllables with which they are associated (Xu 1999), to be discussed in more detail in section 5.3. In the case of particles, the peak delay that may result from the lowering of an L toned mora preceding a particle will occur due to the increased pitch difference between the L and the H targets. Because of articulatory constraints (pitch rises are relatively more difficult to produce than no pitch change or pitch falls according to Xu and Sun 2002) and temporal constraints (the high pitch target needs to be reached quickly in order to align the H tone to the correct mora following the lowered L-toned mora), peak delay has a higher chance of occurring as the pitch distance between the lowered L and the following H increases. In some cases, I argue that this peak delay will “push off” the H tone associated with the particle. I posit that the eventual tonal subordination of formerly H-toned particles was due to this loosening of H’s association with the particle in peak delay situations. In addition, the prevalence of two peaks is relatively low in the first place (18.8% of two and three mora nouns with particles together + 11.2% of LLH converted to HLH + 1.7% of LLF converted to HLF = 31.6%).

Accordingly, I posit that, together, these two factors led to a preference for culminativity, facilitating the tonal subordination of particles and leading to the change of HLH and HLF to HLL. Because of the INITIAL-H constraint, it is the second H tone in these sequences that is deleted in these cases, not the first. In some cases, HLH and HLF will become HHH and HHL respectively instead. This is also due to culminativity, as the changes have reduced the number of separate H tones in each form to a single multilinked H tone. Modeling this stage in Optimality Theory with CULMINATIVITY high-ranked and with the free ranking of MAX-TONE-H and MAX-TONE-L provides an explanation for this variation.

5.2.2 Culminativity in Optimality Theory

The accentual system of Modern Kyoto does not allow accentual units to have more than one prominence (Kubozono 2012), where a prominence is here defined as an H tone, thus respecting the culminativity restriction. This restriction, as I argue in 5.2.1

above, originated in the already relatively low frequency of HLH in two mora nouns with particles and three mora nouns, which eventually decreases even further as particles become tonally subordinate to the nouns they attach to. Formalized here as the constraint CULMINATIVITY, this constraint must be ranked higher than MAX-TONE-H or MAX-TONE-L, depending on which ranking between the two MAX-TONE constraints is chosen by a speaker at evaluation time.

The introduction of this constraint will cause HLH, the output of (49) above, to change into the initial accented HLL. In addition, examination of Sugito (1995) reveals that some LLH patterns also became HHH in Modern Kyoto. As mentioned in the previous subsection, I propose that MAX-TONE-H and MAX-TONE-L are freely rankable. In Stage 1, such free rankability would have had no consequences, as the winning candidate always performed better on both constraints, violating neither. However, in Stage 2, different rankings of the two constraints will produce different outputs, demonstrating their free ranking. I propose that this free ranking accounts for the HLL/HHH divergence in the development pathway.

These developments are shown in the tableaux in (64) and (65) below. Note that I am here assuming that tones which may become adjacent because of the deletion of a tone fuse into a single tone which, despite no longer being two separate tones, will still fulfill the correspondence requirements necessary to satisfy IDENT-IO(TONE). Thus, the deletion of L in HLH will allow one H tone to spread, causing the two original H tones to become adjacent and fuse.

(64) HLH > HLL (MAX-TONE-L >> MAX-TONE-H)

/$H_1L_2H_3$/ *ábùrá* 'oil'	CULM	OCP	MAX-L	MAX-H	INIT-H	NC-μ	ID-TONE	DEP-TONE	NML-H
☞ a. $H_1L_{2,3}$ *ábùrà*				* (H_3)			* ($μ_3$)		
b. $H_1L_2H_3$ *ábùrá*	*! W			L			L		
c. $L_{1,2}H_3$ *àbùrá*				* (H_1)	*! W		* ($μ_1$)		
d. $H_{1,2,3}$ *ábúrá*			*! W (L_2)	L			* ($μ_2$)		** W
e. $H_1H_2H_3$ *ábúrá*	*! W	*!* W		L			* ($μ_2$)		

(65) HLH > HHH (MAX-TONE-H >> MAX-TONE-L)

/$H_1L_2H_3$/ *ágùrá* 'sitting cross-legged'	CULM	OCP	MAX-H	MAX-L	INIT-H	NC-μ	ID-TONE	DEP-TONE	NML-H
☞ a. $H_{1,2,3}$ *ágúrá*				* (L_2)			* (μ_2)		**
b. $H_1L_{2,3}$ *ágùrà*			*! W (H_3)	L			* (μ_3)		L
c. $H_1L_2H_3$ *ágùrá*	*! W			L			L		L
d. $L_{1,2}H_3$ *àgùrá*			*! W (H_1)	L	* W		* (μ_1)		L
e. $H_1H_2H_3$ *ágúrá*	*! W	*!* W		L			* (μ_2)		L

As these tableaux show, when MAX-TONE-L is ranked above MAX-TONE-H, HLL results. When their ranking is reversed, HHH results instead. In either case, the ranking of CULMINATIVITY above MAX-TONE-H rules out the fully faithful candidate, HLH. In (64), note that it is the second H which is deleted due to culminativity, resulting in HLL (64a), not the first H, whose deletion would result in LLH (64c). This is due to INITIAL-H.

The tableaux in (67) and (68) below consider HLF, the output of initial H gain on LLF. Like HLH, culminativity changes HLF to HLL. Additionally, examination of words that were originally LLF in EMJ included in Sugito (1995) shows that a portion of these words also become HHL instead of HLL. I posit that in the latter case, HLF become HHF before ultimately becoming HHL. Both developments to HLL and HHL are considered here.

This input is more interesting than the input HLH discussed in (64) and (65) above because it involves one additional tone, the low tone in the falling contour tone associated with the third mora. Modern Kyoto transparently treats a falling contour tone as a sequence of two tones. That this is the case can be seen when a word ending in a falling contour is suffixed with a particle. This causes the falling contour to be split into an H tone on the final mora of a word and an L tone on the following particle, as shown in (66) below.

(66) F is a sequence of H and L in Modern Kyoto
àmê 'rain' → *àmé-gà* 'rain (nom.)'
LF → LH-L

In order to make clear the fact that F consists of two separate tones, F is represented as [HL] in the tableau below. The development from HLF to HLL and HHF is shown in (67) and (68) below.

(67) HLF > HLL (MAX-TONE-L >> MAX-TONE-H)

/$H_1L_2[HL]_3$/ *támàkî* ‘elbow guard’	CULM	OCP	MAX-L	MAX-H	INIT-H	NC-μ	ID-TONE	DEP-TONE	NML-H
☞ a. $H_1L_{2,3}$ *támàkì*				* (H_3)			* (μ_3)		
b. $H_1L_2[HL]_3$ *támàkî*	*! W			L		* W	L		
c. $H_1L_2H_3$ *támàkí*	*! W		*! W (L_3)	L			* (μ_3)		
d. $L_{1,2}[HL]_3$ *tàmàkî*				* (H_1)	*! W	*! W	* (μ_1)		
e. $H_{1,2,3}L_3$ *támákî*			*! W (L_2)	L		* W	* (μ_2)		** W
f. $H_{1,2}L_3$ *támákì*			*! W (L_2)	* W (H_3)			** (μ_2,μ_3)		* W
g. $H_1H_2[HL]_3$ *támákî*	*! W	*!* W	*! W (L_2)	L		* W	* (μ_2)		
h. $H_{1,2,3}$ *támákí*			*!* W (L_2, L_3)	L			** W (μ_2,μ_3)		** W

(68) HLF > HHF (MAX-TONE-H >> MAX-TONE-L)

$/H_1L_2[HL]_3/$ *nánàê* ‘seven layers’	CULM	OCP	MAX-H	MAX-L	INIT-H	NC-μ	ID-TONE	DEP-TONE	NML-H
☞ a. $H_{1,2,3}L_3$ *nánáê*				* (L_2)		*	* (μ_1)		**
b. $H_1L_{2,3}$ *nánàè*			*! W (H_3)	L		L	* (μ_3)		L
c. $H_1L_2[HL]_3$ *nánàê*	*! W			L		*	L		L
d. $H_1L_2H_3$ *nánàé*	*! W			* (L_3)		L	* (μ_3)		L
e. $L_{1,2}[HL]_3$ *nànàê*			*! W (H_1)	L	* W	*	* (μ_1)		L
f. $H_{1,2}L_3$ *nánáè*			*! W (H_3)	* W (L_2)		L	** W (μ_1,μ_3)		* L
g. $H_1H_2[HL]_3$ *nánáê*	*! W	*!* W		* W (L_2)		*	* (μ_2)		L
h. $H_{1,2,3}$ *nánáé*				**! W (L_2, L_3)		L	** W (μ_2,μ_3)		** W

CULMINATIVITY rules out candidates (67b, c, g) and (68c, d, g), all of which have more than one separate H tone. As in (64) and (65) above, the ranking of MAX-TONE-H and MAX-TONE-L will result in different winners. When MAX-TONE-L dominates MAX-TONE-H, L_2 will be preserved and the creation of an H plateau is avoided, resulting in HLL being selected as the winner. As was the case in (64), INITIAL-H will rule out the candidate (67d) without an initial H, LLF. When MAX-TONE-H dominates MAX-TONE-L, the H tone of the contour is preserved in favor of preservation of L_2, resulting in an H plateau and selecting HHF as the winner.

When HHF is processed as a new input by the grammar, however, it will undergo a second change, simplifying its contour. I posit that the contour simplifies to L rather than H in order to preserve both input tones, whereas a change to H would result in the deletion of the L tone. This is shown in (69). Because there is no interaction between MAX-TONE-H and MAX-TONE-L in this tableau, and any ranking of the two constraints will have no effect on which candidate wins, I leave their ranking undetermined in the contest shown in (69), indicated by the dashed line in between them. It is still assumed, however, that a proper ranking between them will be selected at evaluation time.

(69) HHF > HHL

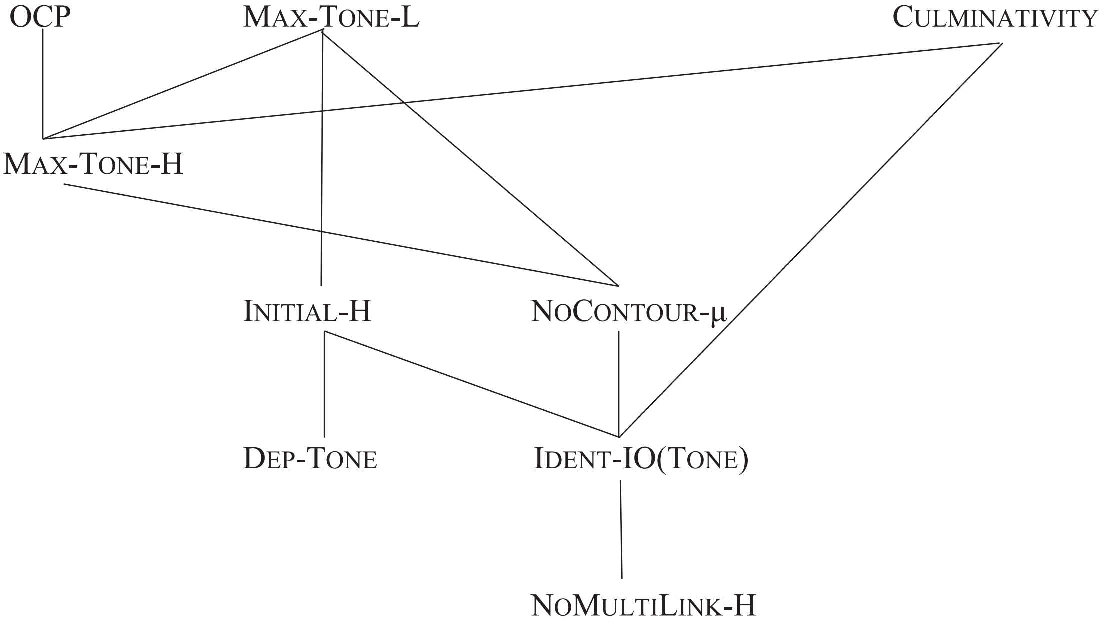
/$H_{1,2,3}L_3$/ *nánáê* ‘seven layers’	Culm	OCP	Max-H	Max-L	Init-H	NC-μ	Id-Tone	Dep-Tone	NML-H
☞ a. $H_{1,2}L_3$ *nánáè*							* (μ_3)		*
b. $H_{1,2,3}L_3$ *nánáê*						*! W	L		** W
c. $H_{1,2,3}$ *nánáé*				*! W			* (μ_3)		** W
d. $H_1L_{2,3}$ *nánàè*							**! W (μ_2,μ_3)		L
e. $H_{1,2}H_3$ (L_3 changed to H_3) *nánáé*	*! W	*! W					* W (μ_3)		* W
f. $H_{1,2}H_3$ (L_3 deleted and H_3 inserted) *nánáé*	*! W	*! W		*! W			* W (μ_3)	* W (H_3)	* W

As the tableau in (69) demonstrates, the contour is simplified due to NOCONTOUR-μ being crucially ranked over IDENT-IO(TONE) and IDENT-IO(TONE) being crucially ranked over NML-H. That NOCONTOUR-μ did not have an effect in selecting HHL over HHF in (68) but *does* have such an effect in (69) is due to the fact that (68) and (69) have different inputs. Changing HL[HL] (i.e. HLF) into HHL in (68) requires deletion of L_2 and H_3, violating both MAX-TONE constraints, whereas changing the input to HH[HL] (i.e. HHF) only involves deletion of L_2, making it the optimal candidate. However, changing HH[HL] to HHL in (69) only requires deleting the association line between the H tone and the third mora, leaving both input tones intact. As such, the MAX-TONE constraints cannot decide between HHL and HHF in (69) as they did in (68), and the decision is left to NOCONTOUR-μ and IDENT-IO(TONE), selecting (69a) HHL over the faithful candidate (69b) HHF.

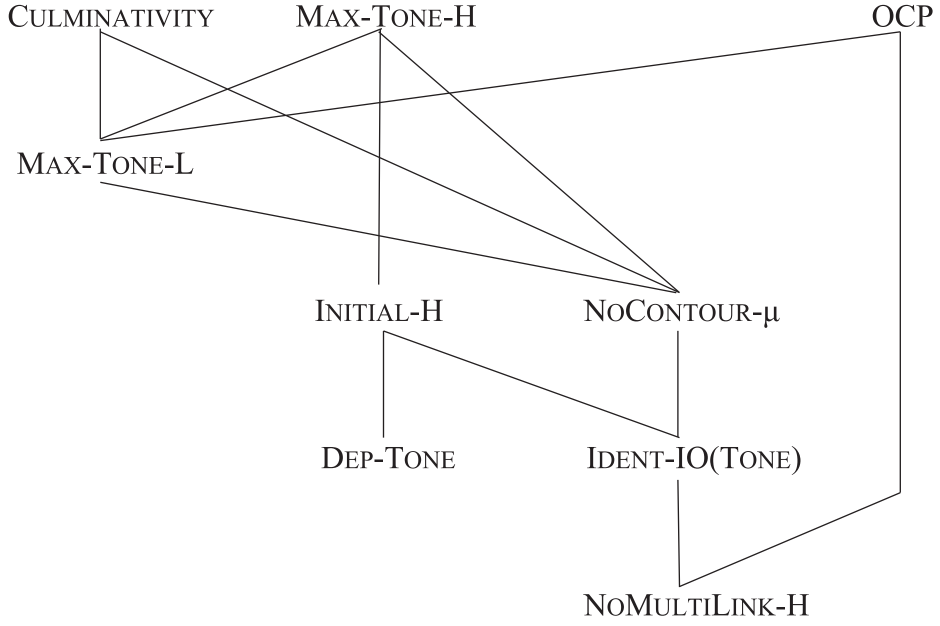
The constraint rankings active in this stage are given in (70) and (71). (70) shows the hierarchy when MAX-TONE-L dominates MAX-TONE-H. Although INITIAL-H showed no mutual interactions with any constraints in this stage, there is no evidence to change its ranking with respect to any constraint except CULMINATIVITY, which was promoted to the highest stratum of the hierarchy. As such, all dominance relations from the previous stage between INITIAL-H and constraints other than CULMINATIVITY are retained in the Hasse diagram below. Note that although there was evidence from candidate (68h) to rank MAX-TONE-L over NO-CONTOUR-μ when MAX-TONE-H dominates MAX-TONE-L (as shown in (71)), there is no evidence to rank MAX-TONE-H over NOCONTOUR-μ when MAX-TONE-L dominates MAX-TONE-

H (as shown in (70)). However, as there is no evidence to change the relation MAX-TONE-H had with NOCONTOUR-μ in the previous stage (MAX-TONE-H >> NOCONTOUR-μ), it is retained here.

(70) Stage 2 Constraint Hierarchy (MAX-TONE-L >> MAX-TONE-H)



(71) Stage 2 Constraint Hierarchy (MAX-TONE-H >> MAX-TONE-L)



As the Hasse diagrams show, the state of both grammars is very similar, with all strata equal with the exception of the variable placement of MAX-TONE-H and MAX-TONE-L in either the first or second stratum. The two grammars differ only slightly with respect to constraint interactions. In the grammar shown in (71), CULMINATIVITY dominates NOCONTOUR-μ, where no such relation could be determined in the grammar shown in (70). In addition, the OBLIGATORY CONTOUR PRINCIPLE can be determined to dominate NOMULTILINK-H in the grammar in (71), but not in the grammar shown in (70).

To summarize, the two strategies taken to avoid multiple H tones in a single word – deletion of the second H, producing HLL, and plateauing between the two Hs, producing HHH and HHL – are explained by two changes to the constraint hierarchy: 1) the reranking of CULMINATIVITY from its relatively low position in Stage 1 to an undominated position in Stage 2, and 2) the free ranking of MAX-TONE-H with respect to MAX-TONE-L in Stage 2, which could not be determined to exist in Stage 1, due to the fact that no competition in Stage 1 crucially depended on a ranking between the two constraints.

5.3 Stage 3: Peak Delay

5.3.1 Rightward Shift of Tonal Associations through Peak Delay

According to Nakai (2001), between the Muromachi era (14th to 16th centuries) and the end of the Edo era (17th to 19th centuries), LHH changed to LLH. Longer patterns like LHHH also changed, yielding LLLH. I posit that these shifts occurred due to peak delay.

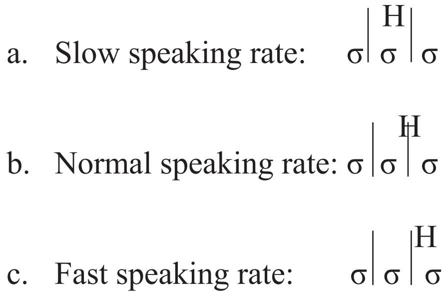
Peak delay is a phenomenon in which an F0 peak occurs later than the syllable with which it is associated (Xu 1999). Peak delay is a well-attested phenomenon and has been reported for a number of languages, including Chichewa (Myers 1998), Mandarin (Xu 1999, 2001), English (Tokuma and Xu 2009), and Japanese (Hata and Hasegawa 1988, Hasegawa and Hata 1992, 1995, Sugito 2001, and Tokuma and Xu 2009).

Based on the results of his production study investigating the relationship between tone (H, weakened H, R), speaking rate (fast, normal, slow), and peak delay in Mandarin, Xu (1999, 2001) reasons that peak delay occurs due to articulatory constraints governing how quickly a speaker can execute a pitch change. He concludes that peaks are more likely to be delayed into a following syllable the less time a speaker has to arrive at a high pitch target within the syllable with which that high pitch target is associated. In addition, Xu and Sun (2002) found that pitch rises require more effort, and thus more time, to execute and complete than pitch falls or no pitch change.

In the previous section, I posited that increased pitch distances between two tonal targets may have the same delaying effect as well. The two situations are parallel. A faster speaking rate results in compressed syllable durations, meaning that less time is available to reach an H target, leading to the need for a pitch rise to be executed more quickly (compared to normal speaking rate syllables) in order for the peak not to be delayed. Similarly, a pitch rise needs to be executed more quickly when there is an increased pitch distance between two tonal targets (compared to non-enhanced pitch distances) in order for the peak not to be delayed. Thus, peak delay occurs due to a combination of the effects of both articulatory and temporal constraints. On an articulatory level, pitch rises require time to execute, and on a temporal level, faster speaking rates and greater pitch distances may result in insufficient time to complete a pitch rise within the correct tone-bearing unit.

Peak delay can be schematized as in (72) below, which shows an H tone associated with the second syllable (the other syllables are left unspecified for the purposes of this illustration) being aligned further and further to the right as the speaking rate increases. Lines between syllables represent syllable boundaries.

(72) Illustration of Peak Delay



Hata and Hasegawa (1988) and Hasegawa and Hata (1992) conducted a production experiment and a perception experiment to test Sugito (1972, as cited in Hata and Hasegawa 1988)’s hypothesis that the percept of initial accent can be saved even with peak delay as long as the delay is compensated for by a steep (fast) pitch fall. They found that when peak delay occurred (24% of the 840 tokens produced) in the production study, pitch falls were steeper, and in the perception experiment, they found that, in general, steeper pitch falls could save the percept of non-delayed accent even with substantial peak delays (up to a threshold, past which even steep pitch falls could not save the percept of non-delay). Based on their results, Hasegawa and Hata conclude that peak delay may serve as a seed for accent shift based on whether a speaker compensates for delay and also depending on which cues a listener pays attention to. If a speaker does not produce a steep enough pitch fall to compensate for peak delay, a listener will end up perceiving an accent that has been shifted to the post-accent syllable. If a speaker *does* produce adequate compensation, then two scenarios may result: 1) the listener may either attend to the compensation and perceive non-delayed accent, or 2) the listener may fail to attend to or otherwise discount the presence of a steep enough pitch fall, perceiving a shifted accent.

I posit that the change from LHH to LLH is due to peak delay, which leads to the H target being shifted rightward onto the final mora. Crucially, this change occurs in the parent generation, not the learning generation. The parent generation will produce peak-delayed forms due to articulatory factors. The learning generation, on the other hand, will learn the LLH input faithfully, converting the peak-delayed outputs of the parent generation into new underlying forms. Furthermore, this rightward shift only affects LHH, but not HHL, HLL, or LHL, resulting in only one change from the Muromachi period to the late Edo period. Why should this be the case? I posit that this is due to a constraint against shifting accent location, NOFLOPACCENT (Alderete 1999). The constraint is defined in (73) below.

(73) NOFLOPACCENT: Do not move accent to another position.

Assign one violation for every accent in the output which has moved from its position in the input.

Thus, the left-hand H of LHH can move rightward to yield LLH because it does not violate NOFLOPACCENT. On the other hand, rightward shift of H in HHL, HLL, or LHL will incur NOFLOPACCENT violations. The existence of such a constraint in a parent generation seems to me to be plausible, as peak delay in Hata and Hasegawa (1988)'s study was observed in only 24% of 840 productions. Though they do not discuss how many tokens exhibited a sufficiently steep pitch fall to compensate for the peak delay and how many did not, it seems likely to me that an even smaller amount than 24% of 840 productions were insufficiently compensated. Even under the assumption that learners consistently ignore compensation by steep pitch fall,[15] at least 76% (in actuality, probably much more) of input data would not show (perceptible) peak delay.

In longer shifts like LHHH > LLLH, I posit that the change occurs in multiple stages through iterated peak delay across learning generations (e.g. LHHH > LLHH > LLLH).

Unlike the previous stages, I argue that the change in the third stage is due to learners producing outputs which are faithful to the input supplied to them by the parent generation. As will be seen in Section 5.3.2, because the change depends on faithfulness and not avoidance of marked structure, no constraints rankings are changed, though several new dominance relations which could not be established in the previous stage can be established here.

5.3.2 Faithfulness to Peak Delayed Inputs in Optimality Theory

In Stage 3, peak delay occurs, pushing the beginning of the H tone in Class 6 LHH rightward, yielding LLH. Crucially, unlike the previous stages, this is a phonetic shift which is found in the output of a subset of the parent generation, not one produced by the learning generation. Thus, the output LLH serves as the input to the learning generation, which will produce LLH faithfully. The change is gradual, and its prevalence will increase over time as more learners learn from peak-delayed inputs

---

[15] Perhaps the default assumption, as learners would not have significant experience with peak delay and compensation for peak delay.

and posit new underlying forms from them. The learning generation will not take an LHH input in this stage and output LLH, as might occur if this change were driven by markedness. Instead, the subset of learners who receive LHH inputs will faithfully produce LHH. Accordingly, this stage is not dependent on avoiding marked structures. Stage 3 simply ensures faithfulness to the input, whereas Stage 1 produced outputs with initial H tones in order to avoid consecutive low tones at the beginning of words, and Stage 2 produced outputs with only one peak in order to avoid having multiple peaks in words.

No new constraints are needed at this stage. This is demonstrated in (74) below for learners who receive the peak-delayed *ùsàgí* ‘rabbit (LLH) from the parent generation with underlying *ùságí* (LHH).

(74) LLH > LLH

/$L_{1,2}H_3$/ *ùsàgí* ‘rabbit’	CULM	OCP	MAX-H	MAX-L	INIT-H	NC-μ	ID-TONE	DEP-TONE	NML-H
☞ a. $L_{1,2}H_3$ *ùsàgí*					*				
b. $L_1H_{2,3}$ *ùságí*					*		*! W (μ_2)		* W
c. $H_1L_2H_3$ *úsàgí*	*! W				L		* W (μ_1)	* W (H_1)	
d. $H_{1,2,3}$ *úságí*				*! W ($L_{1,2}$)	L		** W (μ_1,μ_2)		** W
e. $H_1L_{2,3}$ *úsàgì*			*! W (H_3)		L		** W (μ_1,μ_3)	* W (H_1)	
f. $H_{1,2}L_3$ *úságì*			*! W (H_3)	*! ($L_{1,2}$)	L		*** W (μ_1,μ_2, μ_3)	** W ($H_{1,2}$,L_3)	* W

Because of the role INITIAL-H played in changing LLH to HLH in an earlier stage, it might be expected that HLH or HLL should result again when learners receive an LLH Input in this stage. Interestingly, however, this is not the case. As the tableau shows, outputs with initial H are ruled out by CULMINATIVITY and MAX-TONE-H/L, whose ranking above INITIAL-H now renders INITIAL-H unproductive, with its effects now being felt only when a decision has to be made between retaining an initial H or a final H (as shown above in example (64) in Stage 2 in Section 5.2). With only (74a) and (74b) remaining, IDENT-IO(TONE) selects the maximally faithful candidate, (74a). The joint effect of these three constraints results in the selection of (74a) as the output on the basis of faithfulness.

As mentioned above, the subset of the parent generation with lower rates of peak delay are more likely to produce LHH as their output, which serves as the input to the learning generation. Learners who are exposed to this output will learn LHH instead of LLH, through faithfulness to the input that they receive. This is confirmed in (75), using non-peak-delayed *ùságí* (LHH) 'rabbit' as the input.

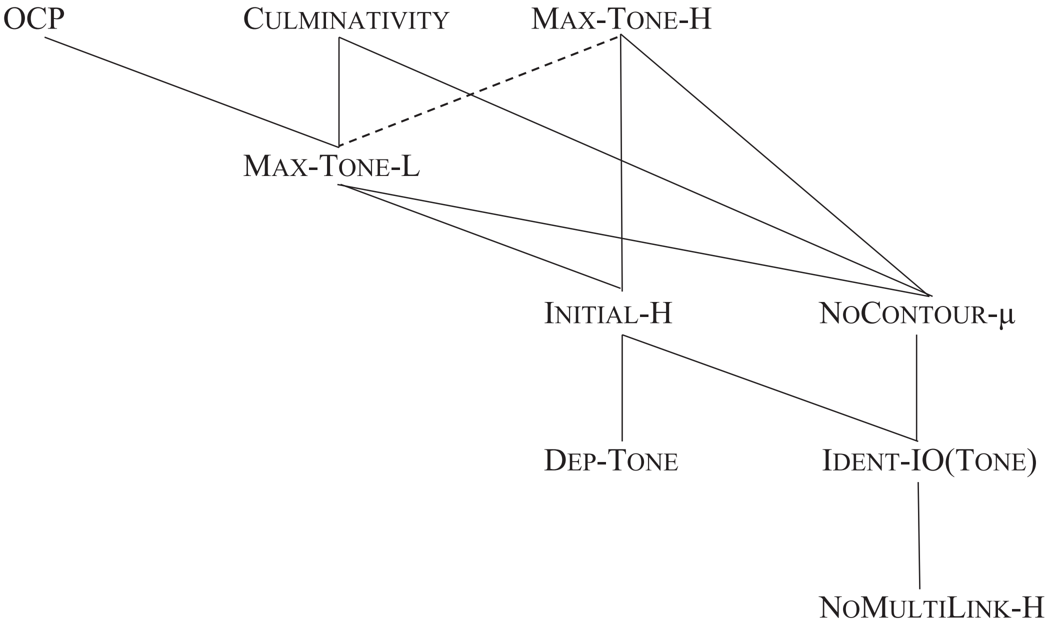
(75) LHH > LHH

/$L_1H_{2,3}$/ *ùságí* ‘rabbit’	CULM	OCP	MAX-H	MAX-L	INIT-H	NC-μ	ID-TONE	DEP-TONE	NML-H
a. $L_{1,2}H_3$ *ùsàgí*					*		*! W (μ_2)		L
☞ b. $L_1H_{2,3}$ *ùságí*					*				*
c. $H_1H_{2,3}$ (L_1 changed to H_1) *úságí*	*! W	*! W			L		* W (μ_1)		*
d. $H_1H_{2,3}$ (L_1 deleted and H_1 inserted) *úságí*	*! W	*! W		*! W (L_1)	L		* W (μ_1)	* W (H_1)	*
e. $H_{1,2,3}$ *úságí*				*! W (L_1)	L		* W (μ_1)		** W
f. $H_1L_{2,3}$ *úsàgì*			*! W (H_3)		L		*** W (μ_1,μ_2, μ_3)	** W (H_1,L_2)	L
g. $H_{1,2}L_3$ *úságì*				*! W (L_1)	L		** W (μ_1,μ_3)	* W (L_3)	* W

As more learners learn from outputs containing LLH, more communities will develop the shifted LLH form from LHH.

Although no constraints have been re-ranked in this stage, several new dominance relations have been established which could not be determined in the previous stage. Both MAX-TONE-H and MAX-TONE-L must dominate INITIAL-H, as shown in (74d-f) and (75d-g), though it is unclear whether the two MAX-TONE constraints must be ranked with respect to each other, as neither of the optimal outputs of (74) and (75) violate either constraint. I leave the two constraints unranked with respect to each other in the tableaux above and the Hasse diagram below. However, as the previous stage demonstrated, CULMINATIVITY and OCP must dominate either MAX-TONE-H or MAX-TONE-L, depending on which grammar is selected in order for CULMINATIVITY and OCP to prefer forms which do not have two peaks. Accordingly, I propose that whichever of the two MAX-TONE constraints is ranked lower is placed on a second stratum which is lower than the stratum including CULMINATIVITY and OCP, but higher than the stratum including INITIAL-H. In addition, candidates (74c) and (75c-d) show that CULMINATIVITY and OCP must be ranked above INITIAL-H. Finally, as in the previous stage, IDENT-IO(TONE) must be ranked over NOMULTILINK-H as shown by (75a). Other rankings are maintained from the previous stage, as there is no evidence to change rankings other than those previously mentioned. The constraint hierarchy at this stage are given in (76) below. The dotted line between MAX-TONE-H and MAX-TONE-L is used here to indicate that these constraints may be reranked with respect to each other with no effect on the other rankings.

(76) Stage 3 Constraint Hierarchy



The last stage of the development of Modern Kyoto is the reduction of multilinked H tones.

5.4 Stage 4: Avoidance of Multiply Linked Hs Where Possible

In the last stage of development, Class 2 HHL and Class 4 HLL ~ HHL change to HLL. This change began in the late Edo period (Nakai 2001) and was recently completed (Frellesvig 2010).

In the previous stage, I argued that LHH became LLH due to peak delay, resulting in fewer multilinked H tones. I posit that this shift is the seed of a change toward avoiding multiply linked H tones wherever possible. I posit that this reduction in the amount of adjacent moras bearing an H tone led the learning generation (or a subpart of it) to promote a constraint that reduces multiply linked H tones. I formalize this in the constraint NOMULTILINK-H, following Ito and Mester (2018)'s usage of the constraint for Kagoshima Japanese.

This constraint will cause HHL to change to HLL or LHL. Which change occurs will be discussed below. Class 1 HHH, which violates NOMULTILINK-H, will remain HHH despite incurring two violations of the constraint. I propose that this is due to the constraint DEP-TONE, which prohibits insertion of a new tone.

5.4.1 Avoidance of Multiply Linked Hs in Optimality Theory

The classes that are affected by the promotion of NOMULTILINK-H are Class 2 HLL and the HHL subset of Class 4 HLL ~ HHL, merging both forms to HLL. NOMULTILINK-H must be promoted and ranked above IDENT-IO(TONE). This is shown in the tableau in (77) below. Note that DEP-TONE has also been promoted to the highest stratum; this will be explained following the tableau.

(77) HHL > HLL

/$H_{1,2}L_3$/ *ázúkì* ‘red bean’	CULM	OCP	DEP-TONE	MAX-H	MAX-L	NML-H	INIT-H	NC-μ	ID-TONE
☞ a. $H_1L_{2,3}$ *ázùkì*									* (μ_2)
b. $H_{1,2}L_3$ *ázúkì*						*! W			L
c. $L_{1,2,3}$ *àzùkì*				*! W ($H_{1,2}$)			* W		** W (μ_1,μ_2)
d. $H_{1,2,3}$ *ázúkí*					*! W (L_3)	*!* W			* (μ_3)
e. $L_1H_2L_3$ *àzúkì*			*! W (L_1)				* W		* (μ_1)
f. $L_{1,2}H_3$ *àzùkí*			*!* W ($L_{1,2}$)	*! W ($H_{1,2}$)	*! W (L_3)		* W		*** W (μ_1,μ_2,μ_3)

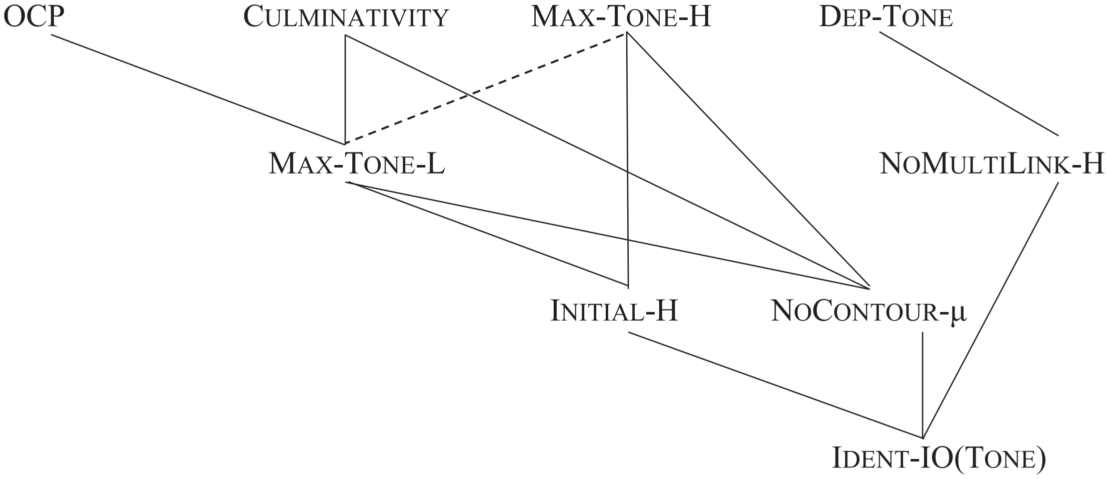
As discussed in previous sections, Class 1 HHH generally remains HHH, which would be ruled out by NOMULTILINK-H in favor of a candidate with fewer association lines with H. This can be explained by the promotion of the constraint DEP-TONE, which must be ranked above NOMULTILINK-H, preventing the insertion of a new tone to satisfy NOMULTILINK-H. This is shown in (78) below.

(78) HHH > HHH

/$H_{1,2,3}$/ *kátátí* 'shape'	Culm	OCP	Dep-Tone	Max-H	Max-L	NML-H	Init-H	NC-μ	Id-Tone
☞ a. $H_{1,2,3}$ *kátátí*						**			
b. $H_{1,2}L_3$ *kátátì*			*! W (L_3)			* L			* W (μ_3)
c. $H_1L_{2,3}$ *kátàtì*			*! W ($L_{2,3}$)			L			** W (μ_2,μ_3)
d. $L_{1,2}H_3$ *kàtàtí*			*! W ($L_{1,2}$)			L	* W		** W (μ_1,μ_2)
e. $L_1H_{2,3}$ *kàtátí*			*! W (L_1)			* L	* W		* W (μ_1)
f. $L_1H_2L_3$ *kàtátì*			*!* W (L_1,L_3)			L	* W		** W (μ_1,μ_3)

The following Hasse diagram in (79) illustrates the state of the grammar in Stage 4. The new rankings established in this stage place Dep-Tone over NoMultiLink-H, which is ranked above Ident-IO(Tone). I assume all other rankings from the previous stage are retained. As above, the dotted line between Max-Tone-H and Max-Tone-L indicates that the two constraints can be reranked with no consequences for the other rankings.

(79) Stage 4 Constraint Hierarchy



5.4.2 On a Possible Pathway from HHL to LHL

A further problem remains: while Nakai (2001) writes only about HHL's change to HLL, and Frellesvig (2010) claims that HHL has become HLL recently, the constraints and rankings above do not account for Frellesvig's note that HHL *can* become LHL and Sugito (1995)'s data which show that some HHL words in EMJ became LHL in Modern Kyoto. Several of these cases seem to be amenable to explanation by invoking the sonority hierarchy (de Lacy 2007) and a requirement that when a choice is available between two moras, and one mora is more sonorant than the other, the prominence should be placed on the more sonorant mora. Words in Sugito (1995) where this seems to be the case include *tùbásà* ‘wing,’ *nìkógè* ‘downy hair,’ *hùtáè* ‘two layers,’ *mìdórì* ‘green,’ and *mùkádè* ‘centipede.’ Not all words can be attributed to this generalization; these include *kàsíwà* ‘oak,’ *òrókà* ‘fool,’ *kìbísù*/*kùbísù* ‘heel,’ *kènúkì* ‘tweezers,’ *tùrúbè* ‘well bucket,’ *hànágì* ‘nose ring for cattle,’ *hògétà* ‘sail,’ and *sìkímì* ‘Japanese star anise.’

However, all of the examples given which follow the sonority generalization have at least three speakers in Sugito (1995) who produce them as LHL. *Tubasa,* and *nikoge* are produced as LHL by three speakers, *hutae* as LHL by four speakers, *mukade* as LHL by five speakers, and *midori* as LHL by all six speakers. Additionally, among the 36 words which were HHL in EMJ included in Sugito’s dictionary, only one, *tóbárí* ‘curtain,’ has the sonority-based ingredients to be pronounced LHL but is not; it is instead pronounced as HHH by all six speakers.

Of those which do not follow the sonority generalization, only the following are produced by at least three speakers as LHL: *kasiwa* (all six speakers), *oroka* (three speakers), and *turube* (three speakers). Of the remaining words, *kibisu/kubisu* and *hanagi* are produced by only one speaker as LHL, and *kenuki*, *hogeta*, and *sikimi* are produced by two speakers as LHL.

One more consideration concerns vowel devoicing and prominence shifting. The words *hutae* ‘two layers’ and *sikimi* ‘Japanese star anise’ have a first vowel which can be affected by vowel devoicing, which can occur when a high vowel occurs between two voiceless consonants. According to Shimabukuro (2007), accent tends to shift rightward in Tokyo Japanese when the originally accented vowel is devoiced. It seems reasonable this may be related to sonority as well. Vowel devoicing would make the vowel less sonorous to the point of being unable to carry pitch information, shifting prominence to the following, more sonorous vowel.

I leave the issue of exactly how HHL could have become LHL in some cases to future research. This future research would need to determine whether a sonority hierarchy-based explanation would be sufficient, investigating the frequency of sonority-based prominence placement both in originally-HHL words and throughout the nominal lexicon.

The proposal that I have presented in Sections 5.1 to 5.4 is an account of language change based on output transmission from a parent generation to a learning generation. In the next section, I briefly describe a model of cultural transmission by Kirby, Dowman, and Griffiths (2007), an abstract discussion of how transmission could work for the development of INITIAL-H, and finally a simple Maximum Entropy-based simulation of the transmission process and development of INITIAL-H using Hayes and Wilson (2008)'s UCLA Phonotactic Learner.

6 The Influence of Cultural Transmission on Language Change

6.1 A Model of Cultural Transmission

According to Kirby, Dowman, and Griffiths (2007), henceforth KDG, the traditional assumption about language form is at its core a biological claim: if there are strong linguistic universals, then this is a reflection of humans' innate ability to acquire and process language and the innate biases and constraints that govern the ability to acquire and process language. Accordingly, if researchers can describe and explain strong linguistic universals, then the collection of these properties can be interpreted as being the structure of the innate human language faculty. However, KDG argue that this traditional assumption fails to account for a crucial factor: the influence of cultural transmission on language evolution. They argue that if the influence of cultural transmission is taken into account, then strong linguistic universals can arise regardless of the strength of prior biases, and thus, strong linguistic universals need not result from strong innate biases and constraints on the acquisition and processing of language.

KDG proposes a model in which the influence of cultural transmission is taken into account along with the influence of biological evolution and the innate biases and constraints on language processing. This model assumes that language is a noisy mapping between meanings and signals, that our innate biases (prior probabilities) favor regularity and systematicity, and that learners are predisposed to choose languages with higher posterior probabilities (probabilities of forms actually observed in input data), even if those languages favor less regularity than learners' innate biases.

The model depends on the concept of iterated learning, in which learners in the current generation receive as input the output of previous generations. Learners will then produce output which serves as the input for a future generation, which will produce output which will serve as the input for a future generation, and so on.

Crucially, learners in the current generation are exposed only to a random subset of the output of the prior generation, not the sum of all outputs that could be generated by all members of the previous generation (as being exposed to such an input would be impossible). The fact that learners are exposed to only a subset of the prior generation's output imposes a "bottleneck" on the information that gets transmitted between groups from generation to generation. The informational bottleneck that a learner receives represents the posterior probabilities of the prior generation's output.

In a test of the model, KDG vary the strength of a prior bias and give the model informational bottlenecks (posterior probabilities) of varying sizes. The results of iterated learning show that even when the strength of a prior bias was weak or zero, the strength of the prior bias had no effect on the distribution of languages predicted by the posterior probabilities of the informational bottlenecks. Thus, the results of KDG's test of the model suggest that strong innate biases are not needed to produce strong linguistic universals, which can instead result from a language distribution characterized by the posterior probabilities of the previous generation. The language a learner learns will then reflect these posterior probabilities rather than whatever language their prior biases might have preferred instead. Thus, cultural transmission may influence and effect language change.

6.2 The KDG Model and Transmission in the Development of Modern Kyoto

To the extent that this model is true, it can be used to account for the phonologization of the increased surface prevalence of H into the constraint INITIAL-H. As discussed in section 5.1 above, a pitch rise enhancement process resulted in an increase in the surface prevalence, which I will now refer to as the posterior probability, of initial H. The posterior probabilities of initial H are now 76.2% (413/542) in bimoraic words and 86.7% (364/420) in trimoraic words. Taking the two together, the overall posterior probability of initial H is 80.8% (777/962). If we assume three groups which have varying levels of initial H posterior probabilities, we can imagine one stage in a transmission process through iterated learning to be like that displayed in (80). Group A displays the baseline level of initial H (230 initial H in bimoraic words + 178 initial H in trimoraic words = 408/962 = 42.4%) with no pitch rise enhancement, Group B displays the increased level of initial H (80.8%) calculated in the discussion in section 4 above with full pitch rise enhancement, and Group C displays an intermediate level of initial H (60% = 577/962) with an intermediate level of application of pitch rise enhancement. In this transmission, I will consider a situation in which each group's informational bottleneck is defined only by the sum of the output of the group from which it descends, with no interaction between groups.

(80) Cultural Transmission of Increased Initial H; Previous Generation to Current Generation (Stage 1, no interaction between groups)

Prev. Gen.:	Group A	Group B	Group C
% Initial H (posterior probabilities):	42.4%	80.8%	60%
Transmission:	↓	↓	↓
Current Gen.:	Group A	Group B	Group C
% Initial H (phonologized):	42.4%	80.8%	60%

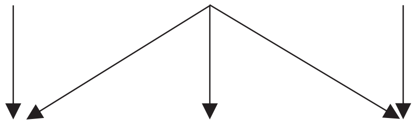
In the previous generation in Group A, no pitch rise enhancement occurred, and so, under the assumption that a learner will choose the language which best reflects the posterior probability of the data that they receive as input, the current learning generation descended from Group A will posit the same frequency of initial H. As a result, they may not have phonologized an initial H constraint or posited a very weak one (as will be seen in Section 6.3, even the baseline level of initial H may contain the seed for the induction of an initial H constraint) and may instead posit the patterns they hear as lexically specified initial H. The previous generation in Group B demonstrated full application of the pitch rise enhancement process, resulting in a high posterior probability of initial H. The current learning generation descended from Group B posits the INITIAL-H constraint due to the high probability of a word having an initial H, phonologizing the posterior probabilities of initial H. Finally, the previous generation of Group C had intermediate implementation of the pitch rise enhancement process. Again, the current learning generation descended from Group C phonologizes what they hear and choose the language with 60% initial H. However, like Group A, their initial H constraint may be weaker than that of Group B, as the application of the pitch rise enhancement process was not total, as it was in Group B. Group C may thus have initial H whose force is felt only in some cases due to relatively low ranking or instead posit the patterns they hear as lexically specified initial H.

In the transmission from the current generation to a future learning generation, let us consider what would happen if there were to be interaction between the groups, as might be expected from contact with migrating/neighboring communities or from language standardization. I will consider a simple interaction in which the current generation of Group B interacts with the next generation of Group A and C so that Groups A and C receive both the posterior probabilities of their parents and the posterior probabilities of Group B. This is shown in (81) below.

(81) Cultural Transmission of Increased Initial H; Current Generation to Future Generation (Stage 2; with interactions between groups)

Current Gen.:	Group A	Group B	Group C
% Initial H (posterior probabilities):	42.4%	80.8%	60%

Transmission:



Future Gen.:	Group A	Group B	Group C
% Initial H (phonologized):	61.6%	80.8%	70.4%

For illustration, the phonologized percentages of initial H in the future generations of Groups A and C are simply the averages of the posterior probabilities received from their parents and from the influential Group B. This interaction results in a snowballing of posterior probabilities of initial H, resulting in greater frequencies of initial H in the future generation. Further interactions between groups and subsequent iterations of cultural transmission will cause further increasing frequencies of initial H until large portions of the community exhibit increased posterior probabilities of initial H, leading to more groups of learning generations to phonologize those posterior probabilities as a relatively strong and influential constraint INITIAL-H.

The next section presents a simple simulation using the UCLA Phonotactic Learner of the transmission schematized in (80), which will result in the eventual phonologization of the INITIAL-H constraint.

6.3 A Simulation of Tonotactic Learning with Maximum Entropy

Hayes and Wilson (2008)'s UCLA Phonotactic Learner (henceforth "the learner") is designed to form constraints on the basis of positive evidence in the input data, a feature set, and the principle of maximum entropy (maxent). Using the principle of maxent, the learner assigns weights to the constraints it finds and assigns probabilities to possible outputs. Hayes and Wilson interpret well-formedness in a maxent grammar as a probability; greater probabilities are judged by the learner to be more acceptable, while lower probabilities are judged to be less acceptable. Probabilities are converted into maxent scores, and forms with higher maxent scores are less acceptable than forms with lower maxent scores. In summary, greater probabilities have lower maxent scores and are judged to be more acceptable, while lower probabilities have higher maxent scores and are judged to be less acceptable.

In this simulation, I show that the UCLA Phonotactic Learner can induce a constraint which can serve as the potential seed for the eventual development of an influential INITIAL-H constraint. The goal of this simulation is to show that, even if the universality of a constraint like INITIAL-H is questionable, it can be induced on a language-specific basis due to the kinds of patterns a learner observes in the input data.

This simulation uses abstract input data in the form of sequences of Hs and Ls with the frequencies varied to simulate the subsets of parent and learning generation pairs shown in (80) above for one generation of transmission. Hs and Ls are defined on the privative features [high] and [low].[16] The feature matrix provided to the learner is given in (82) below. Empty cells are treated as unspecified. Forms with contour tones are currently not included in these simulations.

(82) Feature matrix for H and L

	[high]	[low]
H	+	
L		+

Constraints in the UCLA Phonotactic Learner are formulated in terms of up to 4 natural classes per constraint. A constraint of the form *[NatClass X][NatClass Y] is thus given in terms of 2 natural classes and says that a member of natural class X should not be followed by a member of natural class Y. In this simulation, there are four natural classes, with the member(s) of each class given in parentheses: [+high] (H), [+low] (L), [+word_boundary] (#), and [-word_boundary] (H,L). In order to ensure that the learner learns constraints which do not simply match forms in the learning data (especially in the three mora word cases), I restricted the learner to finding bigram constraints, which consist of only two natural classes.

---

[16] This is primarily for readability. When provided with the binary features [+/-high] and [+/-low], the UCLA Phonotactic Learner uses only [+/-low] in its constraints, treating [+/-high] as a redundant feature and not using it. Since [high] and [low] are clearer than [+low] and [-low], I use privative features in the simulations.

The learners in these simulations are intended to represent different subsets of learners of the same generation who receive inputs from different subsets of parents in the previous generation. These inputs are characterized by different levels of application of the pitch rise enhancement process. The first simulation assumes no pitch rise enhancement (like Group A in (80)), the second simulation assumes an intermediate level of pitch rise enhancement (like Group C in (80)), and the third simulation assumes full pitch rise enhancement occurring wherever possible (like Group B in (80)).

In the first simulation, I provided the learner with the frequencies observed in EMJ as recorded in the *Ruijumyogisho* as input, all multiplied by 100 in order to give the learner enough data to consider. I also provided the learner with accent patterns with particles, as these will create forms which are not observed in the lexical words in EMJ, such as HLH. Recall that particles at this stage have their own independent H tone. These data have the same frequencies as the data without particles. The learning data is given in (83) below.

(83) Learning data (Baseline Initial H; n=37,820)

Without particles		With particles	
H L	1160	H L H	1160
L H	1610	L H H	1610
H H	3040	H H H	3040
L L	3380	L L H	3380
H H H	3170	H H H H	3170

L L L	3000	L L L H	3000
L L H	1120	L L H H	1120
H H L	830	H H L H	830
L H H	790	L H H H	790
L H L	600	L H L H	600
H L L	210	H L L H	210

Given this data, the features above, and the restriction to bigram constraints, the learner induces the following constraints and constraint weights. Constraints are presented in the order they were learned by the learner from first to last.

(84) Constraints and constraint weights (Baseline Initial H)

a.	*[+high][+low]	1.889
b.	*[+low][+word_boundary]	0.301
c.	*[+word_boundary][+low]	0.47

The learner induces the *HL constraint in (84a) and gives it the highest weight. This is transparently due to the fact that patterns containing the sequence HL make up a relatively small proportion of the data, 14.8% (5600/37,820). The constraint *L# in (84b) is induced due to the inclusion of forms with H-toned particles, leaving only 9180 (24.3%) data with word-final L. The most important constraint here is the constraint in (84c), *#L. Although weak in this parent-learner generation pair, with a weight of only 0.47, given the initial input frequencies, the fact that the learner induced this constraint suggests that the accentual pattern frequencies in EMJ in the

Ruijumyogisho already contain the seed for the development of a strong constraint like INITIAL-H.

A test with some testing data was then conducted to see what scores would be assigned to the forms given in the testing data. In this test, no form beginning with L received the perfect maxent score of 0 due to the constraint in (84c), which is to be expected if there is a pressure, even a weak one, to have words with initial H tones. These forms are shown in (85) below. Recall that higher maxent scores are judged less acceptable than lower scores. The following table is ordered in terms of increasing probability/acceptability.

(85) Maxent scores assigned to L-beginning words with Baseline Initial H

Pattern	Maxent Score
LHL	2.66
LHLH	2.359
LL	0.771
LLL	0.771
LH	0.47
LLH	0.47
LHH	0.47
LLLH	0.47
LLHH	0.47
LHHH	0.47

The next simulation tests a group of learners learning from an input showing an intermediate level of increased initial H due to the pitch-rise enhancement process, like Group C in (80). In this simulation, 75% of words which can undergo the pitch-rise enhancement process do. The learning data for this simulation are presented in (86). Note that some forms appear twice with different frequencies (e.g., HHL has

2250 data points in one appearance and 830 in another appearance). This is due to some of the data points resulting from the application of the pitch rise enhancement process. The first of each pair to appear in the list represents pitch rise enhanced forms. The second of each pair represents original forms which did not come about because of the pitch rise enhancement.

(86) Learning data (Intermediate level of initial H, n=37,820)

Without particles		With particles	
H H H	3170	H H H H	3170
L L L	750	L L L H	750
H H L	2250	H H L H	2250
L L H	280	L L H H	280
H L H	840	H L H H	840
H H L	830	H H L H	830
L H H	790	L H H H	790
L H L	600	L H L H	600
H L L	210	H L L H	210
L L	845	L L H	845
H L	2535	H L H	2535
H H	3040	H H H	3040
L H	1610	L H H	1610
H L	1160	H L H	1160

The data that were actually presented to the learner were collapsed so that there were no duplicate forms with different frequencies. The induced constraints are shown in (87).

(87) Constraints and constraint weights (Intermediate level of initial H)

a.	*[+low][+low]	0.688
b.	*[+low][+word_boundary]	0.652
c.	*[+word_boundary][+low]	0.896
d.	*[+high][+low]	0.429

Because of the increase in forms with HL due to the increase in initial H, the *HL constraint (87d) is much weaker in this parent-learner generation pair. The learner also learns a constraint against LL sequences (87a). The constraints in (87b) through (87d) were present in the first generation's grammar, but their weights have increased. On top of that, the INITIAL-H-like *#L constraint *[+word_boundary][+low] (87c) has nearly doubled in strength with a weight of 0.896 compared to the previous simulation's weight of 0.47. Indeed, this is the highest weighted constraint in the grammar induced by this learner. The combined effect of the *LL constraint and the *#L constraint will assign lower probabilities to forms which begin with LL, precisely the context in which INITIAL-H will favor the insertion of an initial H. The effect of this on assigning greater probability to forms which have initial HL than to forms which have initial LL can be seen in the following comparison in (88). The more probable candidate in each horizontal pair, each beginning with an H tone, is bolded.

(88) Initial H preferred to initial L

Form	Maxent Score	Form	Maxent Score
HLL	**1.769**	LLL	2.923
HLH	**0.429**	LLH	1.583
HHL	**1.081**	LHL	1.976
HL	**1.081**	LL	2.236

In the final simulation, I tested a learner which has been exposed to the maximum amount of initial H (initial H frequency of 80.8%, as in Group B in the (80)). The learning data has completely removed LLL, LLH, and LL, adding their frequencies to HHL, HLH, and HL. The learning data is provided in (89) below.

(89) Learning data (full initial H gain, n=37,820)

Without particles		With particles	
H H H	3170	H H H H	3170
H H L	3000	H H L H	3000
H L H	1120	H L H H	1120
H H L	830	H H L H	830
L H H	790	L H H H	790
L H L	600	L H L H	600
H L L	210	H L L H	210
H L	3380	H L H	3380
H H	3040	L H H	1610
L H	1610	H H H	3040
H L	1160	H L H	1160

Again, although there are duplicates in the data above due to some forms resulting from pitch rise enhancement (e.g. HHL with a frequency of 3000), forms which were originally present (e.g. HHL with a frequency of 830) and forms which result from the affixation of a particle (e.g. HLH with a frequency of 3380), duplicate data were consolidated when presented to the learner. The grammar learned by this generation is presented in (90).

(90) Constraints and constraint weights (full initial H gain)

a.	*[+low][+low]	3.186
b.	*[+word_boundary][+low]	1.214
c.	*[+low][+word_boundary]	0.698

The *#L constraint is once again learned by the learner with a weight of 1.214 in (90b), which is stronger than the *#L constraint learned in the previous simulation. The learner also learns the *LL constraint (90a), which is now much stronger when full pitch-rise enhancement has occurred than in the simulation where an intermediate level had occurred. These two constraints together can produce a gang effect on words beginning with LL, assigning greater probability to patterns beginning with HL instead. For example, LLL receives the maxent score 8.285, giving it a lower probability than the corresponding H-gained (via pitch rise enhancement) form HHL, which receives the more probable maxent score 0.698, and the corresponding H-gained (via INITIAL-H) form HLL, which receives the higher, but still more probable maxent score 3.884.

This is a rough simulation of learning with a maxent-based learner in a single iteration of generational transmission to test whether a learner can find a constraint (*[+word_boundary][+low]) and/or set of constraints (*[+word_boundary][+low] and *[+low][+low]) which resemble INITIAL-H given varying levels of frequencies of word-initial H, including the base frequency from the *Ruijumyogisho* without the pitch rise enhancement process. If phonotactic and tonotactic learning can proceed in the way Hayes and Wilson (2008) propose following the principle of maximum entropy, then based on these simulations, it appears that a learner can indeed learn constraints with the effect of INITIAL-H, and it can be suggested that such a language-specific constraint can be induced on the basis of input data.

Naturally, because this is a rough simulation, not all factors have been accounted for, and the present simulation considers only a single iteration of generational transmission rather than several iterations of learning. The simulations as presented would need significant refinement and the implementation of iterated learning in order to track the total development of EMJ to Modern Kyoto, even continuing from the second stage of development (introduction of culminativity), as the grammar of the second simulation above assigns HLH a relatively low maxent score of 0.429, and the third simulation above assigns HLH a perfect maxent score of 0, despite Stage 2 being dependent on some amount of ill-formedness in words having two prominences. This may suggest that an input state like the intermediate level of pitch-rise enhancement given in the second simulation above may be closer to the actual state of the input to learning generations than the maximum occurrence of pitch-rise enhancement given in the third simulation. Further investigation into the usefulness of the maxent learner in modeling language change as well as the construction of more refined simulations should be done in future work.

7 Conclusion

The present paper explored the development of the accentual system of Kyoto Japanese from Early Middle Japanese to Modern Kyoto Japanese. The development of the grammar is accounted for within Optimality Theory, allowing for this development to be captured in terms of competing pressures as constraints whose changing rankings had far-reaching consequences for the grammar and future outputs even in later stages of the development, after a constraint was promoted to higher strata. Motivation for the constraint rerankings proposed in the present account comes from relatively high or low frequencies of certain patterns in data learners would receive as input.

This account focused on the increase in the prevalence of the initial accent pattern, which was relatively uncommon in Early Modern Japanese but which is now the most common pattern in short accented words in Modern Kyoto. This increase in prevalence, along with the development of the system in general, is accounted for here as being the result of four stages of development. In Stage 1, the surface prevalence of initial H is increased due to a pitch rise enhancement process. The increased surface prevalence of initial H is then phonologized by learners in newer generations as a requirement that words have an initial H wherever possible, formalized in the constraint INITIAL-H. In Stage 2, a restriction on the number of prominences in a word (CULMINATIVITY) is promoted to the highest stratum of the grammar. The promotion of CULMINATIVITY in Stage 2 is due to relatively low frequencies of words with two separate peaks (H tones). Most of the changes to initial accent occur as a result of these two stages. In Stage 3, peak delay occurs, changing LHH to LLH, and learners simply need to be faithful to the peak-delayed inputs of their parents. No constraint rerankings were posited in Stage 3, as the developments of Stage 3 were driven by faithfulness, not avoidance of marked structure. Finally, in Stage 4, a constraint against multiply linked H tones (NOMULTILINK-H) becomes active. The promotion of NOMULTILINK-H in Stage 4 is due to the effect of peak delay in Stage 3 reducing the frequency of words with multiple moras linked to H tones. As a result of this promotion, HHL changes to HLL, further increasing the prevalence of initial accent.

In the account presented here, I proposed that learners can induce a language-specific constraint INITIAL-H on the basis of input data which favor patterns such as having a word-initial H. Accordingly, the developmental path in this account uses both a language-specific constraint and constraints thought to be universal, providing an argument for the existence and utility of language-specific constraints in addition to universal constraints.

Also considered in this paper are the conditions under which groups of learners will posit language-specific constraints such as INITIAL-H. It is proposed, based on Kirby, Dowman, and Griffiths (2007), that learners prioritize higher posterior probabilities (or surface prevalence) and posit grammars that will produce languages that match these posterior probabilities. These conditions are necessarily mediated by cultural transmission. A simulation of cultural transmission with different posterior probabilities in Hayes and Wilson (2008)'s maxent-based UCLA Phonotactic Learner suggests that the original frequencies in EMJ already contained the seeds of language change, allowing a learner to induce a weak constraint against word-initial L tones on the basis of input data. As the prevalence of initial H increases, future generations of learners may induce more and more strongly weighted *#L constraints and potentially a *LL constraint which together result in an effect like that of INITIAL-H. Thus, this simulation also suggests that a maxent-based learner like the UCLA Phonotactic Learner can be used fruitfully to uncover the potential seeds of language change, both in Japanese and beyond.

Several issues remain for future research. First, what accounts for the development of HHL to LHL from EMJ to MKJ? One possibility was discussed in section 5.4.2 above, where I speculated that this change may have resulted due to prominence being attracted to more sonorant syllable nuclei. A deeper investigation of accent patterns throughout the lexicon both within and beyond HHL could shed light on this matter.

Second, while I accounted for the increase in initial H and initial accent in the present proposal, a question remains regarding low beginning unaccented words, which account for about 30% of three mora nouns in Modern Kyoto. If LLL words changed to HLL and HHL over the course of Modern Kyoto's development, how did membership in the unaccented LLH pattern increase again? One possible origin for the increased membership may lie in the same mechanism that I propose led to Modern Kyoto LLH in the first place: peak delay, this time applied to the pattern LHL. More changes are likely to be involved, however, as LHL nouns are relatively few (6% of EMJ nouns in Sugito 1995).

Third, while this analysis focuses on trends in the development of Modern Kyoto, why is the development of accent patterns not fully uniform? For example, although Nakai (2001) argues that HHL developed into HLL, and Frellesvig (2010) notes additional correspondences between EMJ HHL and Modern Kyoto LHL, several words in Sugito (1995) which were HHL in EMJ nonetheless show neither pattern in Modern Kyoto, often becoming unaccented, e.g. *tobari* 'curtain,' which is HHH in all six speakers in Sugito's dictionary, or *takigi* 'firewood,' which is produced as HHH by five speakers and LLH by one speaker. The present account attributes such variation to the free ranking of MAX-TONE-H and MAX-TONE-L, but the question remains why one grammar might be chosen over another and whether one grammar is more frequent than the other (and if so, which one?). The answer to this question may be related to the answer to the previous question and may also involve investigating the actuation problem in Japanese, a problem which seeks the answer to the following question: why do specific changes happen at a specific time in a specific place? This third question thus concerns not only variation among speakers of Modern Kyoto but also variation among Japanese dialects together and the question of variation among languages in general.

Finally, how do learners know when high percentages of a certain pattern should be posited to be governed by rules (or constraints) rather than as information to be lexicalized? The Tolerance Principle of Yang (2018) may be able to provide some insight on this question. Relatedly, given the fact that trimoraic nouns in Modern Kyoto are majority unaccented, is it possible for learners to posit rules to account for data which constitute a substantial minority, as in Kubozono (2006)'s proposal that antepenultimate accent is assigned by rule even in native nouns? Work on German plurals (such as Marcus, Brinkmann, Clahsen, Wiese, and Pinker (1995)), in which a minority pattern in plural formation is most productively applied pattern to novel words, may provide some insight as well.

8 References

Alderete, John. (1999). *Morphologically Governed Accent in Optimality Theory*. Amherst, MA: University of Massachusetts Amherst dissertation.

Beckman, Mary E. (1986). *Stress and Non-Stress Accent*. Riverton, NJ: Foris Publications U.S.A.

Campbell, Lyle. (2001). *Historical Linguistics: An introduction* (3rd ed.). Cambridge, MA: The MIT Press.

de Boer, Elizabeth M. (2017). Universals of tone rules and diachronic change in Japanese. *Journal of Asian and African Studies*, 94.

de Lacy, Paul. 2007. The interaction of tone, sonority, and prosodic structure. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, ch.12 (pp. 281-307).

Frellesvig, Bjarke. (2010). *A History of the Japanese Language*. New York: Cambridge University Press.

Goldsmith, John. (1976). *Autosegmental phonology*. Cambridge, MA: MIT dissertation.

Halle, Morris and Jean-Roger Vergnaud (1987). *An Essay on Stress*. Cambridge, MA: MIT Press.

Haraguchi, Shosuke. (1979). Tonology of Japanese Dialects. In George Bedell, Eichi Kobayashi, Masatake Muraki (eds.) *Explorations in Linguistics: Papers in Honor of Kazuko Inoue*. pp. 125-146. Tokyo: Kenkyusha.

Haraguchi, Shosuke. (1999). Accent. In Naoko Tsujimura (ed.) *The Handbook of Japanese Linguistics*, chapter 1, pp. 1-30. Blackwell Publishers.

Hasegawa, Yoko and Kazue Hata. (1992). Fundamental frequency as an acoustic cue to accent perception. *Language and Speech*, 35(1, 2), pp. 87-98.

Hasegawa, Yoko and Kazue Hata. (1995). The Function of F0-Peak Delay in Japanese. *Proceedings of the Twenty-First Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Historical Issues in Sociolinguistics/Social Issues in Historical Linguistics*, pp. 141-151.

Hata, Kazue and Yoko Hasegawa. (1988). Delayed Pitch Fall Phenomenon in Japanese. *Proceedings of the Western Conference On Linguistics WECOL'88*, Vol. 1, pp. 88-100. California State University, Fresno: Department of Linguistics.

Hayes, Bruce and Colin Wilson. (2008). A Maximum Entropy Model of Phonotactics and Phonotactic Learning. *Linguistic Inquiry*, 39:3, pp. 379-440.

Hualde, José Ignacio, Gorka Elordieta, Iñaki Gaminde, and Rajka Smiljanić. (2002). From pitch-accent to stress-accent in Basque. In Carlos Gussenhoven and Natasha Warner (Eds.) *Laboratory Phonology, vol. 7*, pp. 547-584. Berlin: De Gruyter Mouton.

Hyman, Larry M. (1981). Tonal accent in Somali. *Studies in African Linguistics*, 12, pp. 169-203.

Hyman, Larry M. (2009). How (not) to do phonological typology: the case of pitch accent. *Language Sciences* 31, pp. 213-238.

Ito, Junko and Armin Mester. (2007). Prosodic adjunction in Japanese compounds. *MIT Working Papers in Linguistics*, 55, pp. 97-111.

Ito, Junko and Armin Mester. (2015). Sino-Japanese phonology. In Haruo Kubozono (ed.) *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology*, chapter 7, pp. 290-312. Berlin: De Gruyter Mouton.

Ito, Junko and Armin Mester. (2016). Unaccentedness in Japanese. *Linguistic Inquiry*, Vol. 47, No. 3, pp. 471-526.

Ito, Junko and Armin Mester. (2018). Pitch accent and tonal alignment in Kagoshima Japanese. *The Linguistic Review 2018*.

Kawahara, Shigeto (2015). The phonology of Japanese accent. In Haruo Kubozono (ed.) *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology*, chapter 11, pp. 445-492. Berlin: De Gruyter Mouton.

Kirby, Simon, Mike Dowman, and Thomas L. Griffiths. (2007). Innateness and Culture in the Evolution of Language. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 104, No. 12, pp. 5241-5245.

Kisseberth, Charles. (2001). Bantu Perspectives on Yukihiko Nakai's "The History of the Kyoto Accent." In Shigeki Kaji (ed.) *Proceedings of the Symposium Cross-Linguistic Studies of Tonal Phenomena, Tonogenesis, Japanese Accentology, and Other Topics*. pp. 145-151. Tokyo University of Foreign Studies: ILCAA.

Kubozono, Haruo. (2006). Where does loanword prosody come from? A case study of Japanese loanword accent. *Lingua* 116, pp. 1140-1170.

Kubozono, Haruo. (2008). Japanese Accent. In Shigeru Miyagawa (ed.). *The Oxford Handbook of Japanese Linguistics*. New York: Oxford University Press.

Kubozono, Haruo. (2012). Varieties of pitch accent systems in Japanese. *Lingua*, 122:13, pp. 1395-1414.

Kubozono, Haruo. (2019). Secondary High Tones in Koshikijima Japanese. *The Linguistic Review 2019*, 36(1), pp. 25-50.

Labrune, Laurence. (2012). *The phonology of Japanese*. Oxford: Oxford University Press.

Liberman, Mark Yoffe. (1975). The Intonational System of English. Cambridge, MA: MIT dissertation.

Liberman, Mark and Alan Prince (1977). On Stress and Linguistic Rhythm. *Linguistic Inquiry*, 8(2), pp. 249-336.

Marcus, Gary F., Ursula Brinkmann, Harald Clahsen, Richard Wiese, and Steven Pinker. (1995). German Inflection: The Exception That Proves the Rule. *Cognitive Psychology* 29, pp. 189-256.

Martin, Samuel E. (1952). Morphophonemics of Standard Colloquial Japanese. *Language*, 28:3, Part 2: Language Dissertation No. 47: Morphophonemics of Standard Colloquial Japanese (Jul. – Sep., 1952), pp. 7+9-115.

Matsumori, Akiko. (1999). Accentual Reconstruction of the Proto-System of Mainland Japanese Dialects. In Christopher Beedham (ed.) Language *and* Parole *in Synchronic and Diachronic Perspective: Selected Proceedings of the XXXIst Annual Meeting of the* Societas Linguistica Europaea*, St. Andrews, 1998*. New York: Pergamon.

Matsumori, Akiko. (2001). Historical Tonology of Japanese Dialects. In Shigeki Kaji (ed.) *Proceedings of the Symposium Cross-Linguistic Studies of Tonal Phenomena, Tonogenesis, Japanese Accentology, and Other Topics*. pp. 93-122. Tokyo University of Foreign Studies: ILCAA.

Matsumori, Akiko. (2008). On the Reconstruction of the Proto-Accentual System of Japanese. In Bjarke Frellesvig and John Whitman (Eds.) *Proto-Japanese: Issues and Prospects*, John Benjamin Publishing Company.

Mei, Tsu-lin. (1970). Tones and Prosody in Middle Chinese and The Origin of The Rising Tone. *Harvard Journal of Asiatic Studies*, 30, pp. 86-110.

Myers, Scott. (1998). Surface underspecification of tone in Chichewa. *Phonology*, 15, pp. 367-391.

Nakai, Yukihiko. (2001). The History of the Kyoto Accent. In Shigeki Kaji (ed.) *Proceedings of the Symposium Cross-Linguistic Studies of Tonal Phenomena, Tonogenesis, Japanese Accentology, and Other Topics*. pp. 129-143. Tokyo University of Foreign Studies: ILCAA.

Nakai, Yukihiko. (2002). 京阪系アクセント辞典 *[Keihan-kei Akusento Jiten; Keihan-type Accent Dictionary]*. Tokyo, Japan: Bensei Publishing.

National Diet Library Digital Collections. 類聚名義抄 *[Ruijumyogisho]*. http://dl.ndl.go.jp/info:ndljp/pid/2609857?tocOpened=1. Accessed August 4, 2019.

Pierrehumbert, Janet B. and Mary E. Beckman. (1988). *Japanese Tone Structure*. Cambridge, MA: The MIT Press.

Prince, Alan S. (1983). Relating to the Grid. *Linguistic Inquiry*, 14(1), pp. 19-100.

Prince, Alan S. and Paul Smolensky (1993). Optimality Theory: Constraint interaction in generative grammar. Ms., Rutgers University, New Brunswick, NJ and University of Colorado, Boulder. Published, Malden, MA: Blackwell (2004).

Ramsey, Samuel Robert. (1979). The Old Kyoto Dialect and The Historical Development of Japanese Accent. *Harvard Journal of Asiatic Studies*, 39:1, pp. 157-175.

Shibatani, Masayoshi. (1990). *The languages of Japan*. New York: Cambridge University Press.

Shimabukuro, Moriyo. (2007). *The Accentual History of the Japanese and Ryukyuan Languages*. Kent, UK: Global Oriental LTD.

Sugito, Miyoko. (1995). *Ōsaka Tōkyō Akusento Onsei Jiten* [Osaka-Tokyo Accent Voice Dictionary]. Tokyo, Japan: Maruzen.

Sugito, Miyoko. (2001). Production and perception of pitch accent and tone in Osaka Japanese. In Shigeki Kaji (ed.) *Proceedings of the Symposium Cross-Linguistic Studies of Tonal Phenomena, Tonogenesis, Japanese Accentology, and Other Topics*. pp. 209-228. Tokyo University of Foreign Studies: ILCAA.

Tanaka, Yu. (2018). Accentuation in Tokyo and Kyoto Japanese: Toward a unified account. The 157th Meeting of the Linguistic Society of Japan.

Tokuma, Shinichi and Yi Xu. (2009). The effect of F0 peak-delay on the L1/L2 perception of English lexical stress. *Interspeech 2009*, Brighton, UK: 1687-1690.

Uwano, Zendo. (1999). Classification of Japanese Accent Systems. In Shigeki Kaji (ed.) *Proceedings of the Symposium Cross-Linguistic Studies of Tonal Phenomena, Tonogenesis, Typology, and Related Topics*, pp. 151-178

Uwano, Zendo. (2012). Three types of accent kernels in Japanese. *Lingua* 122, pp. 1415-1440.

Vovin, Alexander. (1990). The Origin of Register in Japanese and the Altaic Theory. In Ho-Min Sohn and John Haig (eds.) *Japanese/Korean Linguistics Vol. 6*. Stanford: Center for the Study of Language and Information.

Xu, Yi. (1999). F_0 peak delay: when, where, and why it occurs. *The 14th International Congress of Phonetic Sciences*, San Francisco. pp. 1881-1884.

Xu, Yi. (2001). Fundamental Frequency Peak Delay in Mandarin. *Phonetica*, 58, pp. 26-52.

Xu, Yi and Xuejing Sun. (2002). Maximum speed of pitch change and how it may relate to speech. *Journal of the Acoustic Society of America*, 111(3), pp. 1399-1413.

Yang, Charles. (2018). A User's Guide to the Tolerance Principle. lingbuzz/004146

Yip, Moira. (2002). *Tone*. New York: Cambridge University Press.

Yoshida, Yuko Z. and Hideki Zamma. (2001). The Accent System of the Kyoto Dialect of Japanese: A Study on Phrasal Patterns and Paradigms. In Jeroen van de Weijer and Tetsuo Nishihara (eds.) *Issues in Japanese Phonology and Morphology*. New York: Mouton de Gruyter.