On the interaction of fortition, lenition and *Rendaku* voicing in Japanese
— Experimental and diachronic insights*

Yoshihisa Kitagawa and Aaron Lee Albin
(Indiana University: kitagawa@iu.edu)
IULC Working Papers, 23.1, 1-70.

Abstract

This study proposes an "underlying *ɸ* analysis" of the so-called *ha-gyoo* '*h*-column' consonants in present-day Japanese as an alternative to McCawley's (1968) "underlying *p* analysis." It is argued that an input *ɸ* may surface as any of *β*, *b*, *p*, *ɸ*, ç and *h* due to the interaction of *Rendaku* voicing, fortition and other assimilatory processes. A series of production experiments reveal that, contrary to the popular belief, both lenition and fortition apply to all voiced obstruents in Japanese, leading them to be predominantly (but not categorically) pronounced as continuants intervocalically and as non-continuants post-nasally. In the pursuit of explanations for these synchronic facts, various diachronic changes in the phonology and orthography of Japanese are also investigated. A wealth of historical evidence, especially in relation to the "confusion of the four *kana* characters," is found to eventually lend support to the proposed analysis of the *ha-gyoo* consonants.

Keywords:
ha-gyoo (は行), *dakuten* (濁点), *han-dakuten* (半濁点), *yotsugana* (四つ仮名), *Gendai Kanazukai* (現代仮名づかい), post-nasal, intervocalic, frication duration, burst, underlying *p*, underlying *ɸ*

---

* Previous versions of this paper were presented at *Workshop on Rendaku Encyclopedia* organized by National Institute for Japanese Language and Linguistics (5/18/2013), *Formal Approaches to Japanese Linguistics 7* (6/27/2014), *Buckeye East Asian Linguistics Forum 2* (10/21/2016), and several other informal occasions. The authors are grateful to Isabelle Darcy, Stuart Davis, Ken de Jong, Mark Irwin, Junko Ito, Richard Janda, Shigeto Kawahara, Armin Mester, Charles Quinn, Shin'ichiro Sano, Tomoaki Takayama, Barbara Vance, Yiwen Zhang, and the participants of the numerous oral presentations of this work for their invaluable comments and pointing out the presence of relevant literature and/or data. The authors would also like to thank the anonymous reviewer of IULC Working Papers for carefully reading the manuscript and pointing out the mistakes. The usual disclaimers apply to whatever mistakes and deficiencies originated from our ignorance.

Contents

1. Introduction:

The current work attempts to illuminate the way *Rendaku voicing* is realized in present day Japanese, examining its interaction with the fortition and lenition phenomena in this language. This examination is pursued from various angles. First, it conducts phonetic experiments involving the actual speakers of Japanese and the analyses of their results (Section 2). Second, an attempt to elucidate the essence and the cause of the findings in the phonetic experiments is made, taking into consideration the correlations between their synchrony and diachrony (Sections 3~4). Third, it investigates the interaction of the sound system and the orthography of the Japanese language (Section 4).[1]

1.1. Traditional solution to a puzzle involving *Rendaku* and a newly proposed analysis:

Rendaku 'sequential muddying/voicing', one of the most extensively researched phonological phenomena in Japanese, involves voicing the initial voiceless obstruent in the right member of a pair of compounded elements. The three word pairs in (1) illustrate the kind of voiceless-voiced alternation created by *Rendaku*.

(1) *Rendaku voicing*:

a.	k→g	**k**ana	(か゛な)	*'kana* script'	→	hira-**g**ana	(ひら**が**な) 'cursive *kana*'
b.	t→d	**t**o	(と)	'door'	→	ama-**d**o	(あま**ど**) 'rain door (=shutter)'
c.	s→z	**s**uʃi	(す゛し)	'sushi'	→	maki-**z**uʃi	(まき**ず**し) 'rolled sushi'

Orthographically, the addition of voicing to an obstruent is indicated by the addition of two short lines ゛ called *dakuten* 'turbid dots' on the *kana* 'syllabary character' for its voiceless counterpart, e.g., か → が in (1a).

Generally speaking, *Rendaku* can be characterized as a straightforward voicing process, i.e., one that does not involve any other sound change (in terms of place or manner of articulation). This holds true for all of the examples in (1). A well-known puzzle, however, is that, when applied to *h*, *ç* or *ɸ*, *Rendaku* derives *b*. This is unexpected because *h→b* and *ç→b* involves a change in place of articulation (glottal/palatal to bilabial) and all three (including *ɸ→b*) involve a change in manner of articulation (fricative to stop). The following word pairs illustrate this phenomenon.

(2) *Rendaku* in *ha-gyoo*:

a.	**h**ana	(は゛な)	'nose'	→	waʃi-**b**ana	(わし**ば**な)	'eagle (=hooked) nose'
b.	**ç**iki	(ひ゛き)	'pulling'	→	ɸuku-**b**iki	(ふく**び**き)	'fortune-pull (=lottery)'
c.	**ɸ**uɾo	(ふ゛ろ)	'bath'	→	awa-**b**uɾo	(あわ**ぶ**ろ)	'bubble bath'
d.	**h**eja	(へ゛や)	'room'	→	ai-**b**eja	(あい**べ**や)	'shared room'
e.	**h**oɾi	(ほ゛り)	'carving'	→	uki-**b**oɾi	(うき**ぼ**り)	'float-carving (=relief)'

The traditional solution to this puzzle in a synchronic analysis of present-day Japanese is to assume that /p/ underlies *h*/*ç*/*ɸ* in all Japanese vocabulary (except

[1] Assuming that the readership of each of these topics may not necessarily overlap, we will present our discussions with non-experts in mind for the sake of multidisciplinary readability, which some readers may find too fundamental and wordy. We apologize in advance for such potential annoyance.

loanwords), hence *Rendaku* applies directly to /p/ and derives *b* (McCawley 1968:77-79, 86-87). The surface realization of /p/ as *h*/*ç*/*ɸ* can be accounted for in multiple ways. McCawley (1968:124) posits a phonological rule converting /p/ into *ɸ* in initial and intervocalic environments, which in turn is converted into *h* except before *u*. (*ç* must then further be derived from *h*.)[2] Alternatively, Ito and Mester (1986:53) and Kubozono (1999:60) assume *h* (rather than *ɸ*) is derived from /p/, which then undergoes allophonic rules to turn into *ɸ* and *ç*. In either case, deriving *b* directly from /p/ makes it possible to maintain *Rendaku* as a straightforward voicing rule.

Both diachronic and synchronic evidence have been cited as motivation for this analysis. Diachronically, positing an underlying /p/ makes sense given the consensus among Japanese philologists that all *h*/*ç*/*ɸ* were pronounced as *p* in ancient Japanese, which was subsequently replaced with *ɸ* in all vocabulary except onomatopoeia (Hoffmann 1868, Ueda 1903, Hashimoto 1928). Synchronically, the /p/ analysis captures the *p* ~ *h*/*ç*/*ɸ* alternation observed in the following present-day Japanese words.

(3) a. go-***ɸ***uN 'five minutes' ~ ip-**p**uN 'one minute'
b. go-***ç***iki 'five animals' ~ ip-**p**iki 'one animal'
c. go-***h***ai 'five cups' ~ ip-**p**ai 'one cup'

The present study challenges this analysis and instead explores an alternative derivation of *b* from the sound that surfaces as *h*/*ç*/*ɸ* in present-day Japanese. More specifically, the proposal is as follows:

(4) a. *ɸ* underlies all of *h*/*ç*/*ɸ* (for native and Sino-Japanese words)
b. *Rendaku voicing* maps *ɸ* to *β*.
c. An independently-motivated fortition process maps *β* to *b* post-nasally.
d. *ɸ* is mapped to *h*, *ç* and *p* in the relevant environments (to be discussed in the historical settings in Sections 3.2.1, 4.4 and 4.5).

The present study offers two pieces of supporting evidence for this set of analyses: (i) the results from a series of two elicited production experiments, as well as (ii) an in-depth examination of various diachronic changes in the phonology and orthography of Japanese. The overall goal is to arrive at a synchronic analysis that is not only in sync with the historical facts but also properly captures the phonetic facts revealed in the experiments.

1.2. Preliminaries:

Preceding the examinations and the discussions of the proposed analyses, some preliminaries are in order. In Section 1.2.1 below, a brief sketch of the basics of Japanese morpho-phonology and orthography is presented together with our basic stance and the warnings to accompany them. In Sections 1.2.2 and 1.2.3, an overview of the past research on the fortis-lenis alternations in Japanese is presented. The examination of these phenomena in Japanese is crucial in supporting the proposed analysis of *Rendaku* application to *ha-gyoo* consonants as in (4), which may eventually derive *b*.

[2] The description of this rule is not consistent between its various mentions in McCawley's dissertation (McCawley 1965:54, 91, 133) and subsequent monograph (McCawley 1968:56, 79, 124). The formulation of the rule presented here appears to be what McCawley postulated.

1.2.1. Morpho-phonology and orthography of Japanese

To begin with, it is important to clarify the present study's assumptions about lexical strata (i.e., classes of vocabulary) in Japanese since they become crucial in the discussion (and, indeed, have already been referenced above to a limited extent). Present-day Japanese is traditionally assumed to consist of at least four strata, based largely on etymological criteria but also with accompanying morpho-phonological characteristics (Crawford 2009:97-113). These will be defined for the purposes of the present paper as follows. First, "native Japanese" words (*wago* or *Yamato-kotoba*) refer to words for which at least one constituent morpheme is of Japanese origin (even if the entire word as a whole is not, since *Rendaku* applies to the native morpheme in question as in the example *dʒuu**bako*** (重箱) 'multi-tiered box'). Second, "Sino-Japanese" words (*kango*) refer to both loanwords from China that became incorporated into Japanese as well as words coined in Japan using Chinese-loaned morphemes (*wasei kango*). Third, "loanwords" (*gairaigo*) refer to *any* foreign word loaned into Japanese – i.e., not just Western loanwords specifically. (When the pronunciation of a Sino-Japanese word at the time of borrowing is at issue, it will be referred to as a "Chinese loanword.") Fourth and finally, "onomatopoeia" refers to the class of mimetic words, representing either sounds/voices (*giongo* /*giseego*) or states (*gitaigo*).

Next, the transcription conventions adopted in the present study are worth clarifying. The five vowels in Japanese will be indicated by the traditional *a*, *i*, *u*, *e*, and *o*, suppressing the details of their pronunciation (e.g., *u* as [ɯ]). Phonologically long vowels (consisting of two moras) will be written as two consecutive vowels (e.g., *aa*). The so-called moraic nasal will be transcribed with a small-cap *ɴ*, without indication of the allophonic variation in its pronunciation. The intention behind the use of IPA symbols for obstruents in the present paper is summarized in the (5) below. (Note that this table is *not* intended to inventory the phonemes of Japanese but rather captures the range of their potential surface realizations.) For the sake of readability, the more traditional symbols *ʃ* and *ʒ* (palato-alveolar) will be used in place of the more phonetically accurate symbols *ɕ* and *ʑ* (alveolopalatal).[3]

(5) IPA symbols for obstruents in Japanese used in this paper

	Bilabial	**Dental**	**Alveolar**	**Alveopalatal**	**Palatal**	**Velar**	**Glottal**
Stop	p / b	t / d				k / g	
Fricative	ɸ / β	ð	s / z	ʃ / ʒ	ç	ɣ	h
Affricate			ts / dz	tʃ / dʒ			

As described by Vance (1987:18, 20, 23), *t*/*d* are considered by many scholars to involve a distinct place of articulation from *s*/*z*/*ts*/*dz*. While the latter are generally regarded as alveolar, all authors that recognize the distinction agree that *t*/*d* are (i) articulated more toward the front of the mouth and (ii) involve the upper teeth as (at least part of) the upper articulator. While there appears to be no consensus as to how this

[3] To mark their status as approximants, a diacritic down-tack is often added to the phonetic symbol for the fricatives [β̞], [ð̞] and [ɣ̞]. For notational simplicity, however, these diacritics will be omitted throughout the present work except in a handful of cases where they serve as a useful reminder of their status as approximants.

articulation should be characterized, *t*/*d* are classified as (apico-)dental here, following Nihon Onsei Gakkai (1976:492), Kawakami (1977:32-33) and Labrune (2012:3.3). The same place of articulation is assumed to hold for *ð*, a surface variant of *d*.

Finally, for brevity's sake, *h*/*ç*/*ɸ* will be referred to collectively as *ha-gyoo* ('*h*-column') consonants throughout this work. This nomenclature reflects the following organization of the Japanese syllabaries (*hiragana* and *katakana*), referred to as *gojuuon* ('the fifty sounds'):

(6) *Gojuuon* ('the fifty sounds') in Japanese

	wa	ɾa	ja	ma	**ha**	na	ta	sa	ka	a
		ɾi		mi	**çi**	ni	tʃi	ʃi	ki	i
N		ɾu	ju	mu	**ɸu**	nu	tsu	su	ku	u
		ɾe		me	**he**	ne	te	se	ke	e
	wo	ɾo	jo	mo	**ho**	no	to	so	ko	o

Following the traditional reading direction of Japanese, the columns of this table are read from right to left (*a*, *i*, *u*, *e*, *o*, *ka*, *ki, ku, ke, ko, ...*). The table is organized into five rows (*dan*), each indicating a different vowel phoneme (*a*/*i*/*u*/*e*/*o*). Each column (*gyoo*), with the exception of the first and last, corresponds to a different consonant phoneme, e.g., *k* for the second column. Each (non-empty) cell in this table corresponds to a single *kana* ('syllabary character') used in present-day Japanese. Note that the transcription in each cell reflects allophonic variants (e.g., /s/+/i/=[ʃi]). The different consonant columns are referred to by the name of the first cell in that row (e.g., *ka-gyoo* for *ka*/*ki*/*ku*/*ke*/*ko*). Hence, the term *ha-gyoo* 'h-column' will be used to refer to the *ha*/*çi*/*ɸu*/*he*/*ho* series of central interest in the present study.

Before beginning the detailed discussions, we would like to clarify our stance on the nature of *Rendaku* voicing. First, it has been debated in the literature whether *Rendaku* should be regarded as a phonological process at all or merely an idiosyncratic property of individual lexical items. The traditional view, e.g., as expressed in Takayama (1992-3:117-8), is that *Rendaku* is a higher-order morphophonological process that evolved out of the phonological assimilation rule that applies in newly-created compound words. On the contrary, based on a review of several experimental studies, Vance (2014) concludes that *Rendaku* is too irregular to pass as a true hard-and-fast 'rule', although, at the same time, it is too productive in newly coined words (and too pervasive in the existing vocabulary) to be regarded as entirely lexical. Vance ultimately leans toward Ohno's (2000) view that *Rendaku* is an analogical process, i.e., speakers apply *Rendaku* out of analogy from the existing vocabulary upon encountering a new vocabulary item. Whether one regards *Rendaku* as a rule-based phonological process or an analogical process, the actual surface realization of *Rendaku* is bound to involve irregularities and exceptions reflecting such factors as lexical idiosyncrasies, dialectal variations and historical changes/accidents. Under both views, it is also considered that native speakers of Japanese do have some knowledge based upon which they produce or hold clear intuitions about mapping of voiceless obstruents to their voiced counterparts. The only goals of the present work are: ***(i)*** to clarify how such native speakers' knowledge about *Rendaku* interacts with fortis-lenis alternations when *ha-gyoo* consonants are involved in present-day Japanese, and ***(ii)*** to inquire into how such a state of affairs has arisen in the history of the Japanese language. To achieve these goals, we will first conduct phonetic

experiments and then historical surveys in the remainder of this paper. We will not attempt, on the other hand, to clarify how exactly exceptions of *Rendaku* voicing arose or to formalize a morphologically-conditioned phonological process of *Rendaku*.

In order to streamline the discussion, *Rendaku* as well as fortition and lenition will be treated as phonological rules (and described in traditional derivational terms). That makes it easier for us to dovetail the synchronic analyses we propose with the relevant diachronic changes and correlate the two in an intuitive way. That, we believe, will also make it easy for those who are familiar with the philological works developed in more traditional terms to follow our discussions.[4]

1.2.2. Introspection-based studies of fortis-lenis alternations:

One well-known observation is that voiced obstruents in Japanese undergo changes in manner of articulation. Manner alternations of this type occur (i) between an affricate (fortis) and a fricative (lenis), as in *dz~z* and *dʒ~ʒ*, and (ii) between a stop (fortis) and an approximant (lenis), as in *b~β*, and *g~ɣ*. Considering that affricates and stops are non-continuant sounds whereas fricatives and approximants are continuants, this alternation can be thought as fundamentally one involving continuancy. Much of the existing literature seems to consider that these alternations arise from the weakening of non-continuant sounds. See, for example, Hattori (1951:184), Arisaka (1959:58) and Kawakami (1977:52-54) for *dz~z* and *dʒ~ʒ*, Kawakami (1977:37) for *g~ɣ*, Komatsu (1981:256) and Kawakami (1977:32) for *b~β*, and Vance (1987:19, 24), Vance (2008:76), and Amanuma, Otsubo and Mizutani (1997=1978:65-78) for all of these alternations.[5] In the following examples, underlying /dz/, /dʒ/, /b/ and /g/ in native words have traditionally been considered to be weakened to z, ʒ, *β* and *ɣ*, respectively.[6]

(7) **Lenition**

a.	/ʃi**dz**uku/	→	[ʃizuku]	'drip'
b.	/ne**dʒ**iɾu/	→	[ne**ʒ**iɾu]	'twist'
c.	/a**b**aɾeɾu/	→	[a**β**aɾeɾu]	'act violently'
d.	/ha**g**eʃii/	→	[ha**ɣ**eʃii]	'furious'

Where discussed in the literature, however, these manner alternations have often been reported to be merely a tendency rather than a requirement. For example, Kawakami (1977:52-54) emphasizes that in intervocalic positions, *dz* and *dʒ* often, but not always, surface as *z* and *ʒ*. Similar qualification is almost always made for the other alternations as well, e.g., Kawakami (1977:37) for *g~ŋ*, Kawakami (1977:32) and Komatsu

---

[4] Some researchers may prefer to reanalyze what we present below in Optimality theoretic terms, which we will not pursue in the present work. See Ito and Mester 2003 and Ito and Mester 2008, for example, for OT-based phonological analyses. See also the compilation of studies on *Rendaku* in Vance and Irwin (2016), especially Kawahara and Zamma (2016) for an overview of previous generative analyses.

[5] While Kawakami (1977: 32, 33, 37) mentions the weakening of *b* and *g*, he does not mention the weakening of *d* in the same way, which Vance (1987) interprets as indicating that *d* does not weaken in the same way as *b* and *g*. Vance (2008:76), however, has changed his view and says that all voiced stops are weakened to fricatives. The experiments reported in the present study below do in fact demonstrate the existence of a *d~ð* alternation, directly paralleling the *b~β* and *g~ɣ* alternations.

[6] The *kana* representations and the romanized Christian records of the words in (7a-b) through the medieval Japanese period listed in *Nihon Kokugo Daijiten* (2007) suggest that they used to be pronounced with /dz/ and /dʒ/.

(1981:256) for *b~β*, and Vance (1987:19, 24) and Amanuma, Otsubo and Mizutani (1997=1978:65-71) for all alternations in question.

Possibly reflecting such variability they detected, these researchers tended to have refrained from explicitly discussing these sound alternations in terms of phonological rules. Nonetheless, at least implicitly, most of them ascribe their observations to the process of lenition. Some also seem to implicitly postulate fortition for the *z ~dz* alternation, e.g., Hattori (1951), Yamaguchi (2007), Maekawa (2010a), Maekawa (2013).

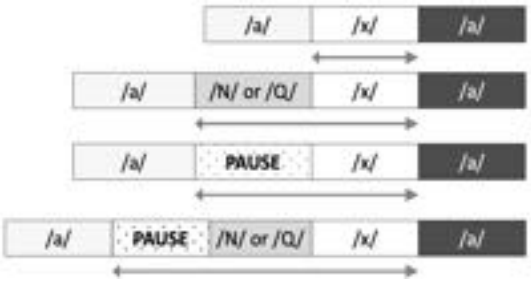
Furthermore, a handful of studies provide additional characterization of the linguistic environments for these alternations. Hattori (1951:184) mentions the realization of *z*/*ʒ* as *dz*/*dʒ* when not immediately preceded by a vowel. Kawakami (1977) presents the surface observation that *dz*/*dʒ* are often realized as *z*/*ʒ*. He then states that this change takes place everywhere in a word except when *dz*/*dʒ* appear word-initially or after a moraic nasal (pp. 52, 54). He also mentions occasional weakening of *b* to *β* in word-internal positions (p. 32). Vance (2008:76) is more specific in identifying the environment for the weakening of /b, d, g/ to [β, ð, γ] as inter-vocalic positions.[7] Yamaguchi (2007:19) also touches upon a positional contrast such that /z/ is allophonically realized as [dz] word-initially and as [z] inter-vocalically.

1.2.3. Corpus-based studies of fortis-lenis alternations:

In most of the studies cited above, the reported insights were based solely on impressionistic judgments. Maekawa (2010a; 2010b; 2013), on the other hand, offers a more formal quantitative study of fortis-lenis alternations in Japanese in a series of works — Maekawa (2010a) for the z/ʒ~dz/dʒ alternation, Maekawa (2010b) for the alternations of b~β, d~ð and g~ɣ, and Maekawa (2013) for all of those alternations. In these works, Maekawa analyzed the phonetic realization of these voiced obstruents in a subset of the Corpus of Spontaneous Japanese. First, based upon the auditory impression of recordings and the visual inspection of waveforms and sound spectrograms, he detected the presence versus absence of articulatory closure in the production of these sounds, appealing to which he distinguished non-continuants (stops/affricates) and continuants (fricatives). He also measured the length of all the obstruents in question as well as their immediately preceding environments, again based upon the auditory and visual inspections. Based upon the analyses of these phonetic measurements, he reported that the rate of articulatory closure for voiced obstruents increases when the Time Alloted for Consonant Articulation (TACA) increases (Maekawa 2010a:369, Maekawa 2010b:8, Maekawa 2013:3037). TACA is defined as the combined durations of the obstruent in question and its immediately preceding non-vocalic segments, as illustrated by Maekawa's Figure 1 reproduced in (8). (Arrows show the durations of TACA. The symbol /x/ stands for voiced obstruents (/z, b, d, g/))

[7] Amanuma, Otsubo and Mizutani's (1977=1978) descriptions of fortis-lenis alternations are rather unclear and appear inconsistent. They seem to assume that *dz* and *dʒ* undergo lenition word-internally before high vowels. (pp. 69, 71) They are, on the other hand, indicating /z/ as an underlying phoneme that are fortified into [dz, dʒ], presumably word-initially and before non-high vowels. (p. 78)

(8) TACA (Maekawa 2013:3037):



Note that TACA is prolonged when the obstruent is immediately preceded by (i) a moraic nasal (*N*), (ii) moraic obstruent serving as the initial consonant in a geminate (*Q*), and/or (iii) a pause (including that which precedes a word/utterance boundary). Observing that voiced obstruents tend to be realized as fortis in more than one position in a word, Maekawa stands for what he calls the coarticulation view of the fortis-lenis variation, and states:[8]

> "Contrary to the traditional linguistic account, the variation is not a positionally conditioned, hence categorical, allophonic variation. … Because speech sounds are produced as the result of the complex coordination of speech organs that are moving simultaneously and continuously, speech production cannot be free from the effect of coarticulation. This is an omnipresent source of variability. Rule-based allophonic variation, on the other hand, is a secondary and optional source of variability." (Maekawa 2010a:360)

Accordingly, he claims that the crucial and essential factor that captures the fortis-lenis variation of a voiced obstruent is the correlation between TACA and the presence versus absence of articulatory closure. Pushing further and extrapolating this view, he conjectures that voiced obstruents in Japanese tend to surface as non-continuants when there is enough time available for them to be articulated with a vocal-tract closure, which is an interesting and reasonable insight into the production of voiced obstruents. All that Maekawa's work offered, however, was the observation on the correlations of surface phonetic phenomena (TACA and closure) and his speculation about what physical aspect of speech production has underlain them. Most importantly, this research does not offer any insight into what compelled the Japanese speakers to alter the articulation of voiced obstruents and perform such linguistic behaviors to begin with. To the contrary, in the current work we are making inquiries into what knowledge of language led Japanese speakers to such linguistic performance. We consider it unlikely (or at least not properly argued) that the speakers fortify or lenite voiced obstruents based upon their knowledge about TACA. Moreover, it is not clear if TACA is entirely free from the notion of phonological contexts. Although Maekawa claims that the fortis-lenis variation is not positionally conditioned but controlled mainly by the influence of TACA, TACA may in fact be regarded as a phonetic artifact that incorporates and reflects the positional information translated into time. In other words, the various types of TACA in (8) above can be regarded as derivatives of such phonological environments as: (i) a position between two vowels, (ii) a position after a moraic consonant (*N* or *Q*), and (iii) a position

[8] Maekawa (2010a:360) describes the term coarticulation as follows: "... in the current study, the term will be used consistently to refer to continuous and noncategorical variations."

after a pause (including that which precedes a word/utterance boundary). TACA, in other words, is the phonetic consequences of a phonological process being significantly controlled by locational factors.

Maekawa obviously presupposes that the input obstruents as phonemes do undergo phonological changes in the fortis-lenis variations in question. He makes statements, for example, such as follows: (Emphasis by the authors of the current work)

> "**/z/ can be realized** as an affricate if the speaker has ample time for the articulation, regardless of the **phoneme**'s position in a word." (Maekawa 2010a:367)

> "**Realization of /z/** in Standard Japanese alternates between affricate [dz] and fricative [z]. Similarly, voiced stops alternate between stops and fricatives, namely, /b/ alternates between [b] and [β], /d/ between [d] and [ð], and /g/ between [g] and [γ] … These alternations are all concerned with the weakening of vocal tract closure articulation" (Maekawa 2013:3037)

He does not, however, clarify much detail of the phonological processes he tacitly adopts. Maekawa (2010b:2) and Maekawa (2013:3037) do not seem to even differentiate lenition and fortition, and regard what he assumes to be the derivation of [dz] from /z/ and that of [β, ð, γ] from /b, d, g/ as basically an identical phenomenon.

Taking such a stance may be considered sufficient and adequate by those whose main concerns are surface phonetic phenomena. We are, on the other hand, interested in elucidating what knowledge of language led Japanese speakers to induce such phonological changes in what particular phonological environments. In the current work in particular, we are aiming to show that some instances of surface [b] in present-day Japanese are derived from the underlying /ɸ/ with the application of *Rendaku* voicing and then fortition (i.e., /ɸ/ → [β] → [b]), as proposed in Section 1.1. It therefore is essential for us to confirm that the Japanese grammar involves a phonological process which fortifies a voiced fricative into a voiced stop to complete this derivation. In particular, it makes a crucial difference to us to confirm whether some instances of surface [b] are a faithful realization of the underlying sound *b* or if it has been derived by the fortition of the input sound *β*.

Maekawa also points out repeatedly that the effect of fortition or lenition discussed above is not entirely categorical but continuous and gradient to some extent. In Section 2 below, we will report and discuss the results of phonetic experiments conducted in a lab, in which many properties of the input stimuli were carefully controlled. As will be seen, some results exhibit reasonably clear and significant effects of phonological processes while others reveal more noncategorical and puzzling variability. In the current work, we take the view that the categorical changes designed by phonological rules are often distorted by the influences from the physiological and physical factors in performance, resulting in the noncategorical phonetic realization of the phonological alternations. Moreover, through the investigation of both spontaneously and artificially induced diachronic changes of the Japanese language in Sections 4.1-4.3, we will arrive at our conjecture that somewhat biased and variable application of the phonological process of lenition came to be induced by such historical changes and has remained in present-day Japanese. We will, in other words, come to hold the view that noncategorical variability of fortition and lenition is partly due to performance but partly due to the subtle changes diachronically materialized in phonological processes.

The present work now turns to its main sections. Their overall structure is as follows. First, Section 2 summarizes the phonetic investigations based upon two phonetic experiments we conducted, which aim to establish baseline data for how the relevant set of sounds are pronounced in present-day Japanese. Section 3 then surveys the interaction of voicing, fortis-lenis alternations and the *h*/*ç*/*ɸ* sounds in Japanese from a historical perspective. Section 4 then ties these two strands of inquiry together in a discussion of diachrony-synchrony interactions. As our investigation of the history of the Japanese language proceeds in Sections 3 and 4, we will see that the diachronic development of the Japanese writing system (including diacritics) provides insight into the phonological aspects of the *ha-gyoo* puzzle.

The ultimate conclusion is that the proposed analysis is better able to account for the phonetic facts in present-day Japanese and straightforwardly reflect the diachronic evolution which gave rise to them. The proposal in (4) should therefore be regarded at least as a viable alternative to McCawley's analysis. One of our goals, in fact, is to demonstrate that the underlying *ɸ* analysis not only inherits the virtues of McCawley's underlying *p* analysis but also eliminates its inconsistencies.[9]

2. Experimental investigations:

This section reports on two phonetic experiments in which the phonological processes involved in the input stimuli are carefully controlled. The goal of the experiments is to get an accurate picture of how voiced obstruents are pronounced in present-day Japanese amid various phonological processes involving fortis-lenis alternations. The results partially confirm some of the observations reported in the previous section but also uncover some new insights, including several rather puzzling phenomena that are ultimately connected to the diachronic changes discussed in Sections 3 and 4. Experiment 1 tests for the two fricative/affricate alternations (*z*~*dz* and *ʒ*~*dʒ*) as well as one approximant-stop alternation (*b*~*β*). Experiment 2 follows up the first experiment by filling in the missing information on the other two approximant-stop alternations: (*d*~*ð* and *g*~*ɣ*). Since the original motivation for the present study relates to the alternation between *β*~*b*, we are particularly interested in finding out how this alternation patterns with the lenis-fortis alternations of other voiced obstruents (both approximant~stop and fricative~affricate). When we observe that fortis-lenis alternations in Japanese take place in all voiced obstruents (including *β*~*b*), our phonological proposal incorporating the fortition of *β* strikes us as a more naturalistic and feasible analysis.

2.1. Experiment 1:

2.1.1. Materials:

The stimulus materials used for the present experiment were all real Japanese words carefully chosen to exhibit certain combinations of properties. First, the target voiced obstruents were at three places of articulation: bilabial (*b*), alveolar (*z*), and alveopalatal (*ʒ*). These target obstruents were crossed with two "positions" (i.e., segmental contexts): either intervocalic position (V_V) or post-nasal position (N_). Finally, to account for the possibility that obstruents whose voicing was derived through *Rendaku* might pattern

[9] In this work, the term "underlying form" is used merely to refer to the input sound of some sound change, either synchronic or diachronic, setting aside the issue of how abstractly or minimally it should be represented.

differently from obstruents with underlying voicing, the stimuli were further crossed with this two-way distinction (referred to here as "derivedness").

The words representing *Rendaku* voicing were compounds in which the second component (whose initial segment is the target obstruent) is either a native or Sino-Japanese word. In all such words, the voicing is derived, i.e., *z* from *s*, *ʒ* from *ʃ*, and *b* from the *ha-gyoo* consonants (*h*, *ç* or *ɸ*).[10]

The words representing underlying voicing were Western loanwords (mostly borrowed from English) — either ordinary nouns (e.g., *menbaa* 'member') or proper names (i.e., names for people/companies/places, such as *Kanzasu* 'Kansas'). The selected words were generally either a single word with no internal structure, e.g., *nanbaa* 'number', or a compound word whose second component is a Western loanword, e.g., *ringo-dʒuusu* 'apple juice'. In either case, the target obstruent falls in a loanword portion of the stimulus word, hence its voicing can be presumed to be underlying. (For example, it is uncontroversial to assume that the voicing of [b] in a word like *basu* 'bus' is underlying, i.e., /b/ rather than derived from /p/.)

Thus, the core of the stimulus materials is structured in a 3×2×2: three places of articulation (bilabial/alveolar/alveopalatal) crossed with two positions (intervocalic/post-nasal) crossed with two kinds of derivedness (*Rendaku*-voiced vs. underlyingly voiced). This structure of the design of the materials is summarized in the following table, where place of articulation is represented by rows, derivedness by columns, and position by the distinction between the first vs. second line within each cell.

(9)

			***Rendaku*-voiced**		**Underlyingly voiced**	
a.	Alveolar	V_V	jae-**z**akura	'double cherry blossoms'	ju**z**aa	'user'
		N_	kan-**z**akura	'winter cherry blossoms'	tan**z**ania	'Tanzania'
b.	Alveo-palatal	V_V	ʃio-**ʒ**ake	'salted salmon'	mane**ʒ**aa	'manager'
		N_	gin-**ʒ**ake	'silver salmon'	ben**ʒ**amin	'Benjamin'
c.	Bilabial	V_V	satsuki-**b**are	'May weather'	ha**b**aa	'harbor'
		N_	nihon-**b**are	'ideal weather'	men**b**aa	'member'

It is important to note that the use of *z*, *ʒ*, and *b* in the transcriptions of the words in (9) above and (10) below and in the rest of the discussion in this section is intended to indicate only the alveolar/alveopalatal/bilabial distinction within the design of the materials, and that no claim is being made about position-modulated manner contrasts, e.g., *z* vs. *dz* or *ʒ* vs. *dʒ*.

This design was further crossed with the vowel following the target obstruent. All five Japanese vowels were used: *a*, *i*, *u*, *e*, and *o*. For example, in the above table, the underlyingly-voiced + intervocalic + bilabial cell was represented not only with words like *haa**b**aa* but also a full paradigm of all five vowels as follows.

[10] We are simply investigating here how the obstruents are phonetically realized at surface when their voicing clearly arose from the application of *Rendaku*. Some irregularities and exceptions observed in *Rendaku* voicing therefore should not raise any concern about the validity of our experiment even for those who are reluctant to regard *Rendaku* as a phonological process.

(10) a. _*a* haa**b***a*a ‘harbor’
b. _*i* taa**b***i*n ‘turbine’
c. _*u* ta**b***u*u ‘taboo’
d. _*e* buɾuu**b***e*rii ‘blueberry’
e. _*o* kaa**b***o*ɴ ‘carbon’

Three words were chosen in this way for each of the five vowels.

Thus, the overall dimensionality of the experimental materials is 3 places-of-articulation × 2 positions × 2 derivedness × 5 vowels × 3 specific words. If all combinations of conditions were represented, this would total to 180 words. However, two classes of cases are missing. First, the sequence of sounds *ʒe* does not occur in native words (at least not in the standard Japanese), hence no *Rendaku*-derived *ʒe* words were included. This created a gap of 3 specific words crossed with 2 positions, i.e., 6 words in total. Secondly, at the level of surface phonetic realization, the sequence *zi* does not occur in any vocabulary stratum of Japanese. This created a gap of 3 specific words crossed with 2 positions crossed with 2 kinds of derivedness, i.e., 12 words in total. Combined, these two gaps lowered the total number of stimulus words by 18, from 180 to 162 (60 words with *b* + 54 words with *ʒ* + 48 words with *z*).

Since these stimulus words are being used for a phonetic experiment, the classification of words into these conditions is based on their surface phonetic realization. Thus, since phonological /zi/ sequences in the native vocabulary are palatalized to [ʒi] (e.g., /kaɴ**z**iɾu/ ‘feel’ as *kaɴ**ʒ**iɾu*), any *Rendaku*-derived *si*→*ʃi*→*ʒi* sequence (e.g., *ton-ʒiɾu* ‘pork soup’, from *ʃiɾu* ‘soup’) is treated as representing the alveopalatal (*ʒ*) condition, not the alveolar (*z*) condition.

To keep interpretation straightforward, all Western loanwords selected to represent *ʒ* (including *ʒi* as in *ʧeɴʒi* ‘change’, *eɴʒiɴ* ‘engine’, *oɾeɴʒi* ‘orange’, *ʧaaʒi* ‘charge’, *abeɾeeʒi* ‘average’, and *dameeʒi* ‘damage’) were English words with [dʒ], not [z]. This means that, while the *z* tokens corresponded to continuant [z] in the English source words (e.g., *juu**z**aa* ‘user’), the *ʒ* tokens corresponded to non-continuant [dʒ]. This asymmetry in the design of the stimuli is mostly unavoidable, as Japanese loanwords coming from English words with [ʒ] and [dz] are rather limited and making it difficult to constitute a full paradigm. Note also that since English lacks phonemic /β/ (and possibly allophonic [β] as well), all English source words for the words in this experiment contain *b*, not *β*.

Finally, to keep the continuancy of the input obstruents straightforward, all native and Sino-Japanese words selected to represent *ʒi/zu* were derived from words with *ʃ/s* (e.g., me+***ʒ***i*ɾi* 'corner of eye' and *suʃi*+***z***u* 'sushi vinegar', from ***ʃ***iɾi 'hips' and ***s***u 'vinegar') and *not* words with *ʧ/ts* (e.g., *hana*+***dʒ***i* 'nose blood' and *mika*+***dz***uki* 'crescent moon', from ***ʧ***i 'blood' and ***ts***uki 'moon'). Likewise, all native words selected to represent *ʒu* were derived from words with *ʃ* (e.g., *ʃi**ʒ**uu* 'beginning to end', from ***ʃ***uu 'end') and *not* words with *ʧ* (e.g., *neɴ*-***dʒ***uu* 'whole year', from ***ʧ***uu 'during'). The full list of the stimulus words used in this experiment is provided as (86) in Appendix A.

2.1.2. Procedure:

Twelve college age (19-24) native speakers of Japanese were recruited from the University of Tokyo to participate in the experiment. All speakers were from the vicinity of the greater Tokyo area — Tokyo itself, Kanagawa, or Saitama. (No speaker from Chiba happened to be included in this experiment.) All participants were naïve to the

purpose of the study.

Three separate lists of the 162 stimulus words were produced, thus creating three "blocks" to the experiment. The order of the 162 stimulus words inside each list was pseudorandomized so as to prevent similar combinations of conditions from occurring in adjacent trials. Each of the three lists used a different pseudorandomization. These three lists were used to create three separate presentation files in the software program *Keynote* (Apple Inc.), with each slide in the presentation containing an individual word from a list in 64-point font centered in the screen. Words were written in normal Japanese orthography, with *furigana* (phonetic guide) included in a few ambiguous cases.

To help ensure each data point would be as statistically independent as possible, efforts were made to minimize cross-trial influence from one token to the next. In particular, it was deemed important to encourage participants to produce each target word as a separate utterance rather than read everything together as a single flowing list of words. Towards this end, the presentation files were designed so that a one-second 150 Hz sine wave 'beep' was played at every transition from one slide to the next. From the perspective of the participant, this functioned as a boundary signaling that one trial has ended and the next has begun.

Participants were seated in a soundproof room and used an iPad to proceed through the slides in the three presentation files at their own pace, with a break between files. To encourage participants to proceed at natural pace, they were instructed to avoid uttering the words too slowly or too carefully. Beeps between adjacent trials were played out loud through the iPad's speakers. Participants' productions were recorded using an IC Recorder (Sony ICD-SX900) set to "manual" attached to a microphone (Sony ECM-CR120) set to "high sensitivity" and "no noise reduction." Recordings were made in stereo at a 22,050 Hz sample rate and 16 bit resolution.

Since each speaker read 3 word lists, this produced 3 repetitions of each word. With 12 speakers and 162 words in each list, the total number of tokens for Experiment 1 was therefore 3×12×162 = 5,832.

2.1.3. Acoustic analysis:

The raw set of recordings consisted of 36 soundfiles, one for each combination of the 12 speakers and the 3 lists. These were then chopped into 5,832 individual soundfiles, one for each individual token. The 5,832 individual soundfiles were then converted from stereo to mono to facilitate data processing. For the purposes of coding the data, a spectrogram was generated with the Spectrogram() function for the R programming language available at https://github.com/usagi5886/dsp, using a Gaussian window and a dynamic range of 50 dB.[11] Along with displaying the spectrogram, the audio of each soundfile was also played, thus allowing judgements to be based on not only visual but also auditory information.

Using these various sources of information, the soundfiles were then coded in two ways. First, all obstruents participating in the fricative-affricate alternation (namely alveolar *z~dz* and alveopalatal *ʒ~dʒ*) were coded for frication duration. This choice of measurement follows Mitani, Kitama and Sato (2006:1603), who show a clear separation between Japanese fricatives and affricates using this dimension (with frication duration

[11] The waveform of each soundfile was not displayed when coding the data. Thus, the spectrogram was the only source of information that was drawn upon in the coding process.

being systematically longer in fricatives and shorter in affricates). To make the measurements, the base R function locator() was used to time-stamp the beginning and end of the frication visible in the spectrogram, thus generating two numbers that could be subtracted to calculate the frication duration. Since a clearly-defined region of frication is generally not visible in the spectrogram for voiced stops/approximants (e.g., *b~β*), this measurement was only applied to *z~dz* and *ʒ~dʒ*, hence only 3,672 of the 5,832 files were coded in this way.

Second, each token was coded for whether a burst (i.e., a brief spurt of high-frequency energy) was visible in the spectrogram at the beginning of the target obstruent. Since a burst is created whenever there is a stoppage of airflow, a greater proportion of bursts indicates a tendency for a given sound to be produced as a non-continuant. The qualitative evaluation of whether a burst is visible in the spectrogram has been used for quantifying manner in several studies, e.g., by Kataoka (2010:331) to distinguish stops vs. fricatives in Northern Paiute (Uto-Aztecan). Judgments were made by the second author, and cases of ambiguity due to the qualitative nature of this analysis were accounted for by using a three-way coding: "yes," "no," and "indeterminate." In order for a token to be classified as "yes," the brief high-intensity energy in the spectrum corresponding to the burst needed to be readily identifiable as distinct from the obstruent's frication (if any) and/or the adjacent vowel. Since bursts can appear in a spectrogram for both stops and affricates alike, this coding could be applied to all target obstruents in the present study, i.e. all 5,832 tokens. For an algorithm-based supplemental analysis conducted to independently confirm the reliability of these manual judgements, see Appendix B.

Thus, frication duration data were produced for tokens with *z* and *ʒ*, and 'burst visible' data was generated for all three places of articulation (*z*/*ʒ*/*b*). These two sets of data were created in two separate rounds of coding. To facilitate comparisons between tokens during the coding process, all tokens corresponding to the same word (i.e., pooling across speaker and list/repetition) were processed together in batches.

For the two fricative-affricate alternations (*z~dz* and *ʒ~dʒ*), fricatives (lenis) are predicted to have a longer frication duration but a lower rate of bursts than affricates (fortis).[12] For the approximant-stop alternation *β~b*, on the other hand, only burst information is collected, and it is predicted that stops (fortis) have a higher rate of bursts than approximants (lenis).

2.1.4. Results:

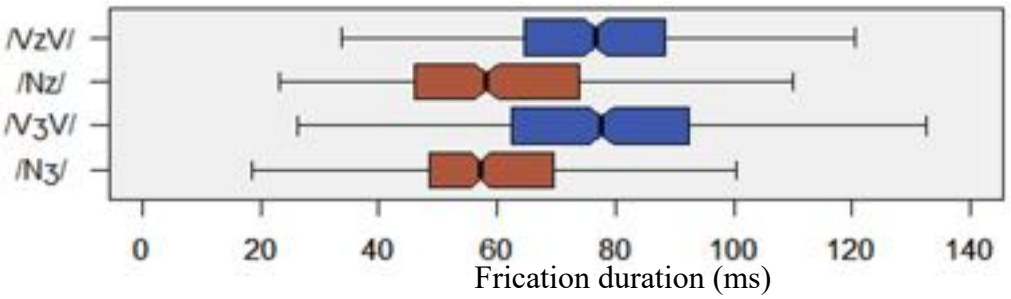
2.1.4.1. Frication duration:

The following boxplot shows the distribution in frication duration for loanwords (i.e., the words with underlying voicing), broken apart into a 2×2 of obstruent and position. The labels along the left side of the plot indicate the input obstruent (*z* vs. *ʒ*) as well as what position it is in (V_V vs. N_). This boxplot represents the distribution of data points for *all measurements* in each combination of conditions, not by-item or by-

[12] In a way, for these sounds, a long frication duration could be thought of as fortis since it would be produced by a closer constriction that is held for a longer period of time (Kingston 2008:20-21). However, for the purposes of the present study, a long frication duration is treated as lenis because it is least affricate-like (cf. Mitani, Kitama and Sato 2006).

subject averages. All other dimensions in the data structure (namely vowel, specific word, speaker, and repetition) have been pooled across in this representation.

(11) Frication duration for loanwords (with underlying voicing)



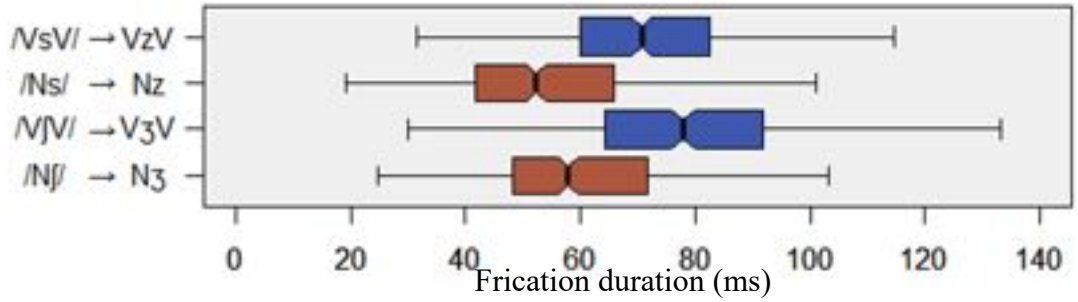
For each condition, the *staple* (thick black vertical bar) indicates the overall median for the relevant set of data points. The *notch* (V-shaped indentation in the box) indicates a sample-based approximation to the confidence interval for the median; thus, non-overlapping notches are a visual hint at what differences are significant. The box indicates the *inter-quartile range* (IQR), i.e., the range of values within which the middle fifty percent (25%-75%) of the data points fall. The whiskers (thin horizontal line with T-shaped caps) indicate the range of the 'core' part of the distribution, outside of which data points may be considered outliers (using 1.5 times the IQR outside the box as a cutoff point). Any outliers outside this range are not shown.

In the plot, it is readily apparent that both obstruents have a systematically longer frication duration in intervocalic position compared to post-nasal position. By using the following R code to compare goodness-of-fit across linear mixed effect models, adding Position (i.e., intervocalic vs. post-nasal) to the model is found to result in a significantly better fit ($p \approx 0.0001$).[13]

(12) anova(lmer(FricationDuration ~ Obstruent + (1|Subject) + (1|Word))),
(lmer(FricationDuration ~ Position + Obstruent + (1|Subject) + (1|Word)))

The following boxplot shows the same kind of data, visually structured in the same way, except that this time the data represents words with *Rendaku*-derived voicing. (Note how the voiceless inputs are reflected in the labels along the left side of the plot.)

(13) Frication duration for words with *Rendaku*-derived voicing



[13] While word length and pitch accent were not controlled for in the design of the stimulus words, post-hoc tests of the same variety as (12), counting word length in moras and coding pitch accent in terms of "Nth mora accented," revealed both had significant effects on the results. However, even with these predictors added into the lmer() formula in (12), the effect of position was still significant. These observations hold true of all analyses presented in this section.

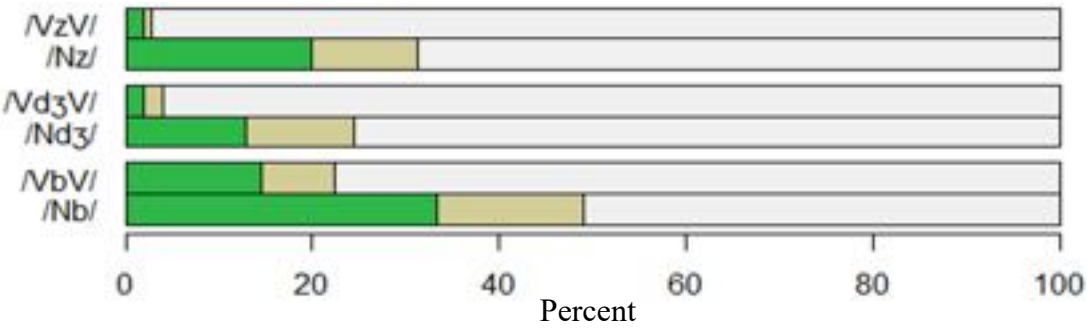
This plot demonstrates how almost identical results obtain for these words as well. Indeed, the goodness-of-fit comparison using the above formula in (12) is also significant here ($p<2.2e^{-16}$). Given the striking resemblance between the plots (11) and (13), it is perhaps not surprising that, if the Derivedness factor is added the model in the second line of (12) (with Position included), the goodness-of-fit does not significantly improve ($p=0.2574$).

Thus, as a whole, the overall conclusion that can be drawn based on the frication duration data is that, irrespective of the distinction between underlying vs. *Rendaku*-derived voicing, the frication of *z* and *ʒ* is of shorter duration post-nasally than intervocalically. This result can be thought of as replicating Mitani, Kitama and Sato's (2006) findings for *voiceless* obstruents (and extending them to *voiced* obstruents as well). In light of Mitani et al.'s finding that shorter frication is a cue to affrication, this indicates that, post-nasally, *z* and *ʒ* are produced in a more affricate-like way, i.e., as *dz* and *dʒ*.

2.1.4.2. Visible bursts:

The results on how frequently bursts were visible in spectrograms for the loanwords (i.e., the words with underlying voicing) are displayed in the following barplot:

(14) Percentage of *loanword* tokens where burst was visible in spectrogram (Experiment 1)



The format of the labels along the left side of the plot is the same as before, with indication of both the obstruent in question as well as the position it appears in. As above, all vowels, specific words, speakers, and repetitions are pooled across in this representation. The bars along the x axis (summing to 100%) represent the percentage of the three different values each token could be coded as. The dark color (green) on the far left indicates "yes" (a burst was unambiguously visible), the intermediate color (beige) in the middle indicates "indeterminate" (it is possible to interpret some aspect of the spectrogram as having a burst), and the lightest color (grey) on the right indicates "no" (there was no burst apparent in the spectrogram). The bars are grouped into pairs according to the intervocalic (V_V) vs. post-nasal (N_) distinction. Comparison of all such pairs of bars makes it clear that there are consistently more bursts post-nasally compared to intervocalically. Note that this hierarchy holds true even for the bilabial (*b*) condition.[14]

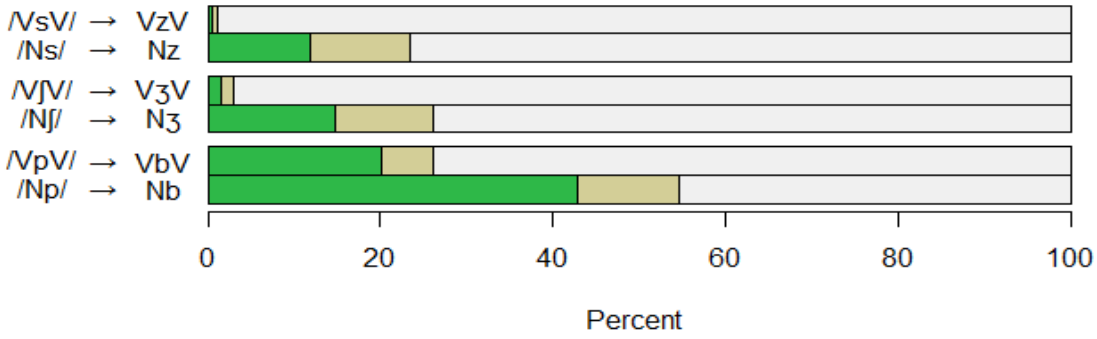
The robustness of this trend was tested by fitting (binomial) generalized linear mixed

[14] The rates of burst appearance for *b*, however, are globally shifted higher (i.e., further to the right) relative to the other two kinds of obstruent (*z* and *ʒ*). This asymmetry will be discussed in Sections 2.3 and 4.3.

effects models, excluding the "indeterminate" tokens and treating the remaining data are as a dichotomous variable (i.e., yes vs. no). Using the same basic structure as in (12) except swapping *lmer(FricationDuration~...)* out for *glmer(BurstVisible~..., family="binomial")*, the goodness-of-fit was found to significantly improve with the addition of Position (i.e., intervocalic vs. post-nasal) as a predictor ($p<2.2e^{-16}$).

The following plot contains parallel data for the words with *Rendaku*-derived voicing, with the voiceless inputs indicated in the labels along the left-hand side of the plot. Note that the labeling for the bilabial (*b*) condition here is indicated in terms of the traditional underlying /p/ analysis of *ha-gyoo* consonants by McCawley (1968) only for the sake of simplicity.

(15) Percentage of *Rendaku* tokens where burst was visible in spectrogram (Experiment 1)



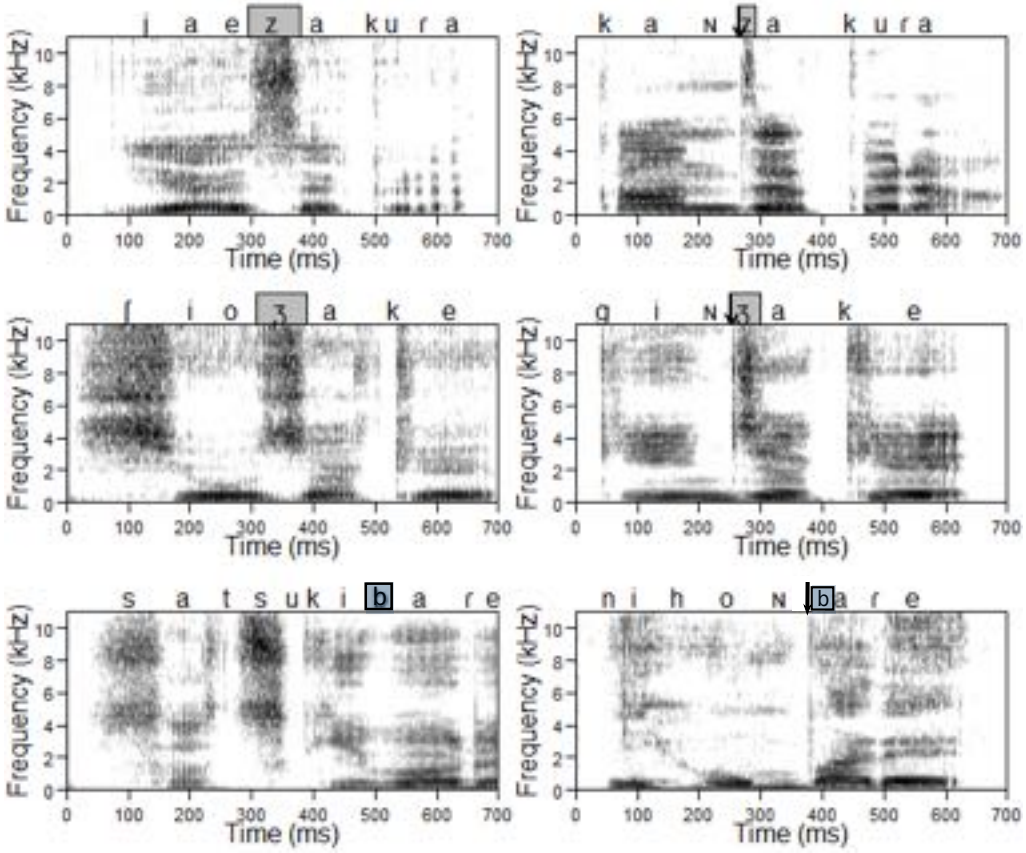
The general patterns here are the same as before: more bursts were visible post-nasally than intervocalically ($p<2.2e^{-16}$). As was found above for frication duration, if the Derivedness factor is added the model (with Position included), the goodness-of-fit does not significantly improve (p=0.5104).

Thus, the overall conclusion is that, irrespective of the distinction between underlying and *Rendaku*-derived voicing, more bursts appear post-nasally. In light of the fact that a burst cues stoppage of airflow, this means that Japanese voiced obstruents tend to surface as fricatives and approximants (i.e., continuants) intervocalically and as affricates and stops (i.e., non-continuants) post-nasally.

2.1.4.3. Summary and illustrative spectrograms:

The findings discussed above for both frication duration and visible bursts are illustrated in the spectrograms in (16) below, which depict representative tokens for *Rendaku*-voiced words. The grey box above each spectrogram indicates the time range of the consonant in question, which roughly corresponds to frication duration for the alternations in first two rows. Cases where the presence of a burst was coded as "yes" are indicated with an arrow placed above the plot in the location where the burst was found.

(16) Spectrograms of typical tokens illustrating the empirical findings from Experiment 1



In the first two rows, frication duration is systematically longer intervocalically (left column) than post-nasally (right column). (Note that the x axis is standardized to [0,700] across all six panels in the plot.) Moreover, bursts are visible only in post-nasal position (as indicated by the vertical grey 'stripe' in the spectrogram under each arrow) and not in intervocalic position. Taken together, these observations suggest these obstruents are produced as fricatives (*z*/*ʒ*) intervocalically and affricates (*dz*/*dʒ*) post-nasally.

In the final row, the token on the right (post-nasal) has a burst, suggesting its status as a stop — i.e., [b]. In contrast, the token on the left (intervocalic) has no burst. Rather, the presence of the consonant is cued by the attenuation of intensity across the entire frequency spectrum (as can be seen in the weakened formant structure). In such cases (which are quite typical in this dataset), the consonant in question appears as the voiced bilabial approximant [β̞]). Note also that there is no region of frication in either of the spectrograms in the final row — precisely the reason why frication duration was not measured for bilabial sounds in this experiment.

These observations suggest that following transcriptions would most accurately capture the production of these sounds in present-day Japanese. Note that the layout of the table below mirrors that of the set of spectrograms.

(17)

	Intervocalic		Post-nasal	
a.	ja*e*-**z***a*kura	'double cherry blossoms'	ka*ɴ*-**dz**akura	'winter cherry blossoms'
b.	*ʃi*o-**ʒ***a*ke	'salted salmon'	gi*ɴ*-**dʒ**ake	'silver salmon'
c.	satsuk*i*-**β***a*re	'May weather'	nihon-**b**are	'ideal weather'

2.1.5. Discussion:

In the following table, whose layout mirrors that of (9) above, the transcription of the example words succinctly summarizes the results of Experiment 1. In parentheses, the assumed derivation of those sounds is provided. Note that the derivation involving underlying *ha-gyoo* consonants is now captured in accordance with the proposed analysis in (4). The purpose of the rest of this section is to show that these assumed derivations constitute a reasonable and viable analysis.

(18)

			***Rendaku*-voiced**	
a.	Alveolar	**V_V** **N_**	ja*e*-**z***a*kuɾa ka*ɴ*-**dz**akuɾa	(s→z) (s→z→dz)
b.	Alveo-palatal	**V_V** **N_**	ʃi*o*-**ʒ***a*ke gi*ɴ*-**dʒ**ake	(ʃ→ʒ) (ʃ→ʒ→dʒ)
c.	Bilabial	**V_V** **N_**	satsuk*i*-**β***a*ɾe niho*ɴ*-**b**aɾe	(ɸ→β) (ɸ→β→b)

First, note that *Rendaku*-derived words like *ka*ɴ*-**dz**akuɾa* and *gi*ɴ*-**dʒ**ake in* (18a-b) (= (9a-b)): (i) start out as voiceless fricatives (*s* and *ʃ*, in *sakuɾa* and *ʃake*), (ii) they are voiced by *Rendaku* to *z* and *ʒ*, and (iii) they are ultimately realized as affricates (*dz* and *dʒ*). This last process can be thought of as a process of *post-nasal fortition*. In intervocalic position, this fortition process does not apply, and the obstruents in question remain fricatives as in *jae-**z**akuɾa* and *ʃio-**ʒ**ake*.

Of central interest for the present paper is the fact that the *Rendaku*-voiced *ha-gyoo* sounds in (18c) can be thought of as undergoing the same kind of fortition process as in (18a) and (18b). If the voiceless input is assumed to be *ɸ* (e.g., /ɸaɾe/), the common assumption that *Rendaku* adds voicing but does not alter place or manner of articulation can be maintained, such that *ɸ* is voiced to *β*. This sound is preserved as-is in intervocalic position (e.g., *satsuki-**β**aɾe*), whereas in post-nasal position, fortition applies and yields surface *b* (e.g., *niho*ɴ*-**b**aɾe*). This analysis applies to all *ha-gyoo* sounds regardless of the following vowel, e.g., *dojoo-**β**i* 'Saturday' vs. *kine*ɴ*-**b**i* 'anniversary' < *ɸi* 'day', *çinoki-**β**uɾo* 'cypress bath' vs. *ɾote*ɴ*-**b**uɾo* 'open-air bath' < *ɸuɾo* 'bath'. The allophonic realization of /ɸ/ as [ç] and [h] can therefore be maintained as an independent process.

The story looks somewhat different for the underlyingly voiced obstruents in the loanwords in (19).

(19)

			Underlyingly voiced	
a.	Alveolar	**V_V** **N_**	j*uu***z***aa* ta*ɴ***dz**ania	(z) (z→dz)
b.	Alveo-palatal	**V_V** **N_**	man*ee***ʒ***aa* be*ɴ***dʒ**amiɴ	(dʒ→ʒ) (dʒ)
c.	Bilabial	**V_V** **N_**	h*aa***β***aa* me*ɴ***b**aa	(b→ β) (b)

Assuming the *underlying manner* of articulation in the Japanese loanword generally matches that of the English source words, English [z, dʒ, b] would be borrowed into

Japanese as /z, dʒ, b/.[15] Thus, English *user* [juzɚ], *Benjamin* [bɛn**dʒ**əmɪn], and *member* [mɛm**b**ɚ] would correspond to Japanese /juu**z**aa/, /beɴ**dʒ**amiɴ/, and /meɴ**b**aa/. Under this assumption, since *z*/*ʒ* in loanwords are assumed to be underlyingly /z/, the same post-nasal fortition process as postulated for the *Rendaku* words can be assumed for loanwords as well. Thus, for example, surface *taɴ**dz**ania* in (19a) can be thought of as the output of post-nasal fortition from underlying /tan**z**ania/ (cf. English [tænzəniə]).

However, the same is not true for the other two obstruents in (19b-c). Out of analogy with English *manager* [mænə**dʒ**ɚ] and *harbor* [hɑɹ**b**ɚ], the underlying forms of the Japanese loanwords may be assumed to have /dʒ/ and /b/, i.e., /manee**dʒ**aa/ and /haa**b**aa/. If so, then the surface forms [man*ee**ʒ**aa*] and [h*aa**β**aa*] must be the output of a lenition process. The results of Experiment 1 suggest such dʒ→ʒ and b→ β mappings only occurs in intervocalic position, hence the relevant process is *intervocalic lenition*.

The full derivations for the example words showing intervocalic lenition are as follows:

(20)				**Lenition**	
a.	(7a)	dz→z	/ʃi**dz**uku/	→	[ʃi**z**uku]
b.	(7b)	dʒ→ʒ	/ne**dʒ**iɾu/	→	[ne**ʒ**iɾu]
c.	(9b)	dʒ→ʒ	/man*ee**dʒ**aa*/	→	[man*ee**ʒ**aa*]
d.	(7c)	b→ β	/a**b**aɾeɾu/	→	[a**β**aɾeɾu]
e.	(9c)	b→ β	/h*aa**b**aa*/	→	[h*aa**β**aa*]

Likewise, the full derivations for the example words showing post-nasal fortition are as follows:

(21)				***Rendaku***		**Fortition**	
a.	(9a)	s→z→dz	/ka*ɴ* + **s**akuɾa/	→	ka*ɴ*-**z**akuɾa	→	[ka*ɴ*-**dz**akuɾa]
b.	(9a)	z →dz	/ta*ɴ***z**ania/	N/A		→	[ta*ɴ***dz**ania]
c.	(9b)	ʃ→ʒ→dʒ	/gi*ɴ* + **ʃ**ake/	→	gi*ɴ*-**ʒ**ake	→	[gi*ɴ*-**dʒ**ake]
d.	(9c)	ɸ→β→b	/niho*ɴ* + **ɸ**aɾe/	→	niho*ɴ*-***β***aɾe	→	[niho*ɴ*-**b**aɾe]

Thus, the experimental results bring to attention something that has left unclear in previous literature on the fortis-lenis alternations. When *Rendaku* voicing is brought into the picture, and if the underlying forms of the words are taken into consideration, in certain cases such alternations are more accurately described as the byproduct of a *fortition* process rather than a lenition process, as illustrated in (21a-c). That is, while many of previous accounts (as summarized in Section 1.2.2) tended to be framed in terms of lenition, there is evidence that lenition and fortition may *both* be at work in the phonology of Japanese. Once fortition is postulated, the possibility arises that it applies generally to *all* voiced continuants (*z*/*ʒ*/*β*) and that (21d) may in fact involve the same kind of derivation as in (21a) and (21c), i.e., *ɸ→β→b*.

2.2. Experiment 2:

Experiment 1 uncovered two classes of intervocalic lenition in loanwords: *dʒ→ʒ* and *b→β*. However, Japanese also has two additional voiced obstruents: *d* and *g*. On one

[15] It is a moot point here whether *z*, *dʒ*, *b* in English were phonemes or phones when they were borrowed into Japanese since these sounds do not show any allophonic changes in these positions in English. We are grateful to Mark Irwin for bringing this issue into our attention.

hand, it is possible that intervocalic lenition is limited to *dʒ→ʒ* and *b→β*, hence *d* and *g* remain unaffected. On the other hand, it is also possible that intervocalic lenition is a general process that applies indiscriminately to all non-continuants, including *d* and *g* (thus creating *d→ð* and *g→ɣ* mappings). The goal of Experiment 2 is to determine which of these two scenarios is the case. Filling in the gaps for these last two sounds completes the picture and helps establish the fuller set of phonetic facts on intervocalic lenition.

2.2.1. Materials:

As with Experiment 1, the stimuli used for Experiment 2 were all real Japanese words structured into various combinations of conditions. This time, the two places of articulation were dental (*d*) and velar (*g*). As before, half of the stimuli had voiced obstruents derived through *Rendaku* and the other half had underlyingly voiced obstruents in Western loanwords.

The most substantial difference between the two experiments is the number of positions that are involved. For the *Rendaku* stimuli, the same two positions were used as in Experiment 1: post-nasal (N_) and intervocalic (V_V). For the loanword stimuli, in addition to these two, word-initial (#_) was included as a third position (e.g., not just compounds with *dansu* 'dance' but also the word *dansu* in isolation). Note that this means the target obstruents in Experiment 2 sometimes occur word-initially and other times word-medially. The fact that *Rendaku* requires a word-medial morpheme boundary as a "trigger" renders word-initial *Rendaku* voicing impossible, hence this third position was only added for the loanword stimuli.

The resulting structure of the stimuli is illustrated in the following table. As before, place of articulation is represented by rows, derivedness by columns, and position by the different rows within each cell.

(22)

		***Rendaku*-voiced**	**Underlyingly voiced**
Dental	**V_V**	gaɾas*u*-**d***a*ma 'glass bead'	ɸook*u*-**d***a*ɴsu 'folk dance'
	N_	ʃabo*ɴ*-**d**ama 'soap bubble'	ɾate*ɴ*-**d**aɴsu 'Latin dance'
	#_	N/A	**#d**aɴsu 'dance'
Velar	**V_V**	ɸuɾ*u*-**g***a*o 'old face'	bod*ii*-**g***aa*do 'body guard'
	N_	ʃi*ɴ*-**g**ao 'new face'	suki*ɴ*-**g**aado 'skin guard'
	#_	N/A	**#g**aado 'guard'

Thus, the overall structure of all of the materials for the second experiment is 2 obstruents (*d* and *g*) crossed with 2 kinds of derivedness (*Rendaku*-voiced vs. underlyingly-voiced) crossed with 2~3 positions (V_V, N_, and possibly #_, depending on derivedness). Moreover, as in the first experiment, this structure was further crossed with 5 kinds of vowel following the target obstruent (*a*/*i*/*u*/*e*/*o*) and 3 specific words.

If all combinations of conditions were represented, this would total to 150 words (2×2×5×3=60 for *Rendaku* words plus 2×3×5×3=90 for loanwords). However, as was the case in Experiment 1, certain classes of cases are missing. More specifically, no words with the sequences *di* and *du* were included in the experiment. Since such sequences are systematically absent from the native and Sino-Japanese vocabulary strata, *Rendaku*-voiced instances of *di* and *du* could not be included in the experiment. This decreased the number of stimuli by 12 (2 vowels × 2 positions × 3 specific words). Since the sequences *di* and *du* do occur in loanwords (e.g., *disuku* 'disc' or *buuduu* 'voodoo'), in principle, it

would have been possible to include such words in the set of stimuli. However, to keep the design simpler and more consistent between the *Rendaku* and loanword conditions, they were excluded. This created a gap of 18 words (2 vowels × 3 positions × 3 specific words). Thus, the total number of stimulus words for Experiment 2 was 150-12-18=120. The full list of the stimulus words used in this experiment is provided as (87) in the Appendix A.

2.2.2. Procedure:

For Experiment 2, ten new participants (age 19-20) were recruited from the same subject pool (at the University of Tokyo) as in Experiment 1. All were from either Tokyo or Saitama. To increase the number of data points, four separate lists of the 120 stimulus words were created (rather than three), thus creating four blocks to the experiment. As before, each list used a different pseudorandomization. The elicitation procedure was the same as in Experiment 1 – using *Keynote* on an iPad to show participants presentation slides with individual stimulus words. As before, sine-wave beeps were inserted at the transitions between slides, and participants were instructed to avoid uttering the words too slowly or too carefully.

Since each speaker read 4 word lists, this produced 4 repetitions of each word. With 10 speakers and 120 words in each list, the total number of tokens for Experiment 2 was therefore 4×10×120 = 4,800 recorded tokens.

2.2.3. Acoustic analysis:

As before, the recordings were chopped into individual soundfiles for each token (4,800 in total) and converted to mono for analysis. Since a subset of the stimuli for Experiment 1 exhibited fricative/affricate alternations, they were coded for frication duration. For Experiment 2, however, the stimuli involve only approximant-stop alternations *d*~*ð* and *g*~*ɣ* (which generally lack frication), hence the same measure was not used. Instead, tokens were only coded in terms of whether a burst was visible in the spectrogram at the beginning of the target obstruent, as described above for Experiment 1.

The same basic procedure was used here: playing the audio of each soundfile, inspecting visualization of the time-varying spectral structure of the audio using the Spectrogram() function, and applying a three-way coding ("yes," "no," and "indeterminate") to classify each token. A total of 4,800 codings were created in this way – one for each of the 4,800 soundfiles. As with Experiment 1, see Appendix B for an algorithm-based supplemental analysis to confirm the reliability of these judgements.

It is well known that, in the mainstream variety of Tokyo Japanese, *g* can also be nasalized to *ŋ* both intervocalically and after the moraic nasal *ɴ*, at varying degrees of consistency depending on the speaker (e.g., Kawakami 1977:34-38, Vance 1987:110-112). Not surprisingly, many of the speakers in the present experiment produced *ŋ*. While coding for the presence vs. absence of a burst alone does not capture this nuance of the data, the environment for nasalization cuts across the two environments under examination (intervocalic and post-nasal). As such, any speakers in the present study exhibiting nasalization presumably would do so (thereby affecting the rates of visible bursts) equally in both contexts. Thus, the existence of nasalization is not expected to compromise the results of the present study. (The fact that *d* and *g* pattern nearly identically in the results below — at least in regard to the contrast between intervocalic

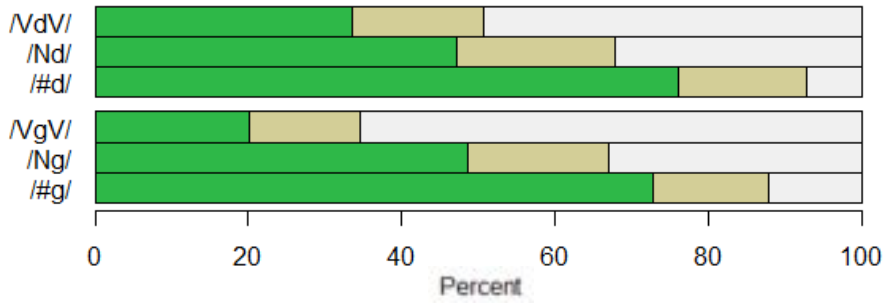
and post-nasal environments — reinforce the plausibility of this explanation.)[16]

2.2.4. Results:

2.2.4.1. Visible bursts:

The results on how frequently bursts were visible in spectrograms for the loanwords are displayed in the following barplot:

(23) Percentage of loanword tokens where burst was visible in spectrogram (Experiment 2)



The format of this plot is the same as the equivalent ones in Section 2.1.4. Here again, the dark color (green) on the far left indicates "yes" (a burst was unambiguously visible), the intermediate color (beige) in the middle indicates "indeterminate" (it is possible to interpret some aspect of the spectrogram as having a burst), and the lightest color (grey) on the right indicates "no" (there was no burst apparent in the spectrogram). Note that the bars are grouped into triplets according to context, i.e., intervocalic (V_V) vs. post-nasal (N_) vs. word-initial (#_).

As was seen for the first experiment, more bursts were visible in the spectrogram post-nasally than intervocalically. Interestingly, the rate of visible bursts is even higher for word-initial position. This is well in accordance with a well-known utterance-initial prosodic "strengthening." We thus take the position that this phenomenon reflects a prosodic effect independent of the phonological contrast involving intervocalic and post-nasal positions.[17] Using the same statistical tests discussed above for the visible burst data in 2.1.4, the goodness-of-fit was found to significantly improve with the addition of

---

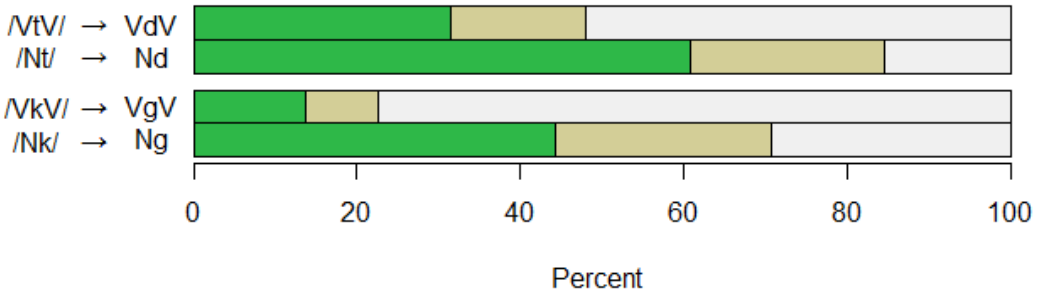
[16] Kamei (1956:12) and McCawley (1968:86-87) argued that the *Rendaku* partner of *k* is (or can be) *ŋ* rather than *g* (for the speakers who use g and *ŋ* distinctively) even in present-day Japanese, taking into consideration the view that *Rendaku* had stemmed from prenasalization of obstruents and the resulting spread of voicing in ancient Japanese. Some present-day dialects are also known to exhibit its residues. (See footnote 19 below for more on this.) Following Vance and Irwin (2016:4) and the contributors in that volume, however, we will refrain from adopting this analysis. As mentioned in the text directly above, *ŋ* and *g* are known to be free variants for many Tokyo speakers. Such free variation can be observed when *k* undergoes *Rendaku* as well. For example, both *oo**ŋ**ama* and *oo**g**ama* 'cauldron (大釜)' are free variants when *kama* 'iron pot (釜)' undergoes *Rendaku*. This phenomenon can be easily captured if *Rendaku* is regarded as a straightforward voicing process to derive *g* from *k* and the nasalization of *g* to *ŋ* is considered to be optional at least for some speakers. If, on the other hand, it is asssumed that *Rendaku* must directly derive *ŋ* from *k* (albeit in more than one step), the free variation in question would remain rather puzzling.

[17] While the pursuit of this topic lies outside the scope of the present study, we cannot help noticing its correlation with Kingston's (2008:1) observation that "[c]onsonants lenite inside prosodic constituents and not at their edges, and lenition therefore conveys to the listener that the current constituent is continuing rather than ending or a new one beginning." Fortis at prosodic boundaries (as suggested by the present results) can be thought of as fulfilling the opposite function at a prosodic edge.

Position (i.e., intervocalic vs. post-nasal vs. word-initial) as a predictor ($p<2.2e^{-16}$). Additional post-hoc tests confirmed this is also true for comparisons of any given pair thereof, namely intervocalic vs. post-nasal, post-nasal vs. word-initial, and intervocalic vs. word-initial ($p<2.2e^{-16}$ for all three).

The following plot contains parallel data for the words with *Rendaku*-derived voicing (Note the voiceless inputs indicated in the labels along the left-hand side of the plot.) Since *Rendaku* does not apply word-initially, we are not taking /#d/ and /#g/ into consideration here, and only two bars (not three) in each group are indicated. The patterns here directly parallel the equivalent bars in (23), with more bursts visible post-nasally than intervocalically ($p<2.2e^{-16}$).

(24) Percentage of *Rendaku* tokens where burst was visible in spectrogram (Experiment 2)



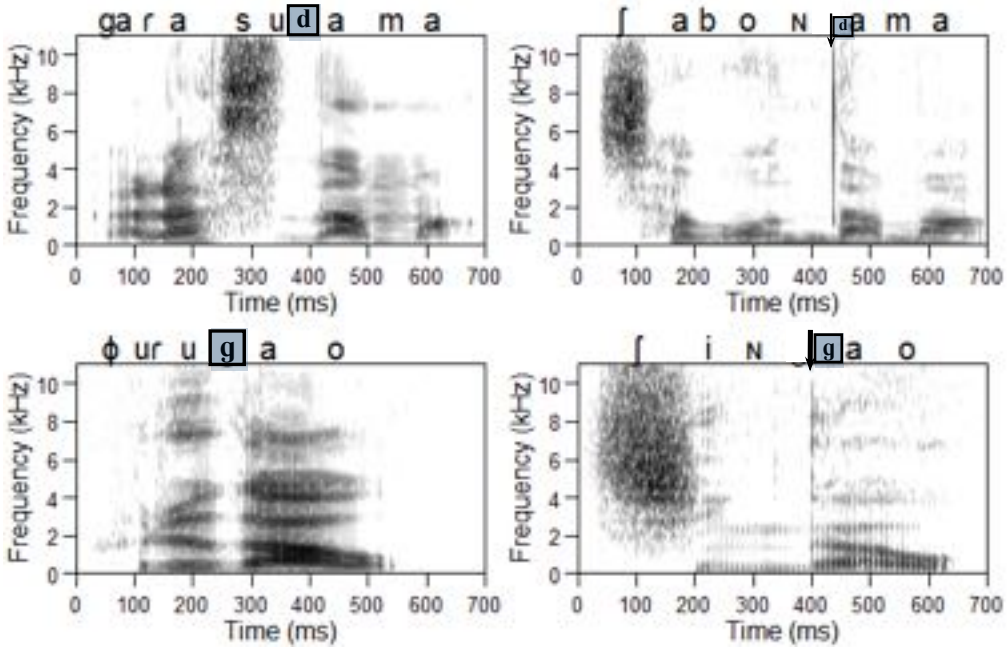
As was found for Experiment 1, if the Derivedness factor is added to the model (with Position included), the goodness-of-fit does not significantly improve. This is true both if word-initial position is included ($p=0.8951$) as well as if it is excluded in order to make the comparison more balanced between *Rendaku* words and loanwords ($p=0.8995$).

Thus, the results of Experiment 2 directly parallel those of Experiment 1: irrespective of the distinction between underlying and *Rendaku*-derived voicing, more bursts appear post-nasally. Thus, the conclusion made earlier (namely, that Japanese voiced obstruents tend to surface as non-continuants post-nasally and as continuants intervocalically) is reinforced.

2.2.4.2. Summary and illustrative spectrograms:

These findings are illustrated in the following spectrograms, which depict representative tokens for *Rendaku*-voiced words from Experiment 2. As before, a grey box indicates the consonant in question, and cases where the presence of a burst was coded as "yes" are indicated with an arrow placed above the plot in the location where the burst was found.

(25) Spectrograms of typical tokens illustrating the empirical findings from Experiment 2



For the post-nasal tokens on the right, bursts are visible (as indicated by the vertical grey "stripe" in the spectrogram under each arrow), suggesting the obstruents in question are stops — i.e., [d, g]. In contrast, the intervocalic tokens on the left have no such bursts. Rather, as was seen earlier for *b*, the presence of the consonant is cued by the attenuation of intensity across the entire frequency spectrum (as can be seen in the weakened formant structure), suggesting they have been lenited to the approximants [ð̞, ɣ̞]).

These observations suggest that the obstruents in question are produced as stops (*d*/*g*) post-nasally and approximants (ð̞/ɣ̞) intervocalically, and as such, that the following transcriptions would most accurately capture the production of these sounds in present-day Japanese. Note that the layout of the table mirrors that of the set of spectrograms.

(26)

	Intervocalic		**Post-nasal**	
a.	gaɾas*u*-**ð̞***a*ma	'glass bead'	ʃabo*ɴ*-**d**ama	'soap bubble'
b.	ɸuɾ*u*-**ɣ̞***a*o	'old face'	ʃi*ɴ*-**g**ao	'new face'

2.2.5. Discussion:

In the table below, the results of Experiment 2 are summarized schematically using the words from (22). As before, the assumed derivation of these sounds (to be discussed immediately below) is shown in parentheses.

(27)

		***Rendaku*-voiced**	**Underlyingly voiced**
Dental	**V_V**	gaɾas*u*-**ð***a*ma (t→d→ð)	ɸook*u*-**ð***a*ɴsu (d→ð)
	N_	ʃabo*ɴ*-**d**ama (t→d)	ɾate*ɴ*-**d***a*ɴsu (d)
	#_	N/A	**#d**aɴsu (d)
Velar	**V_V**	ɸuɾ*u*-**ɣ***a*o (k→g→ɣ)	bod*ii*-**ɣ***aa*do (g→ɣ)
	N_	ʃi*ɴ*-**g**ao (k→g)	suki*ɴ*-**g**aado (g)
	#_	N/A	**#g**aado (g)

First, note that in words with *Rendaku*-derived voicing, the obstruents in question start out as voiceless stops (*t* and *k*, in *tama* and *kao*), which *Rendaku* changes to voiced stops *d* and *g*. These obstruents cease to be stops and change into the approximants *ð* and *ɣ* in intervocalic position. This process can be thought of as the same kind of lenition observed in Experiment 1 for *b*~*β* and *dʒ*~*ʒ* in loanwords. In post-nasal position, this lenition process does not apply, hence the obstruents remain stops. The full derivation of the lenited words is therefore as in (28a-b), in contrast to the derivation for the non-lenited words in (28c-d):

(28)

			Rendaku		**Lenition**	
a.	t→d→ð	/gaɾas*u* + **t***a*ma/	→	gaɾas*u*-**d***a*ma	→	[gaɾas*u*-**ð***a*ma]
b.	k→g→ɣ	/ɸuɾ*u* + **k***a*o/	→	ɸuɾ*u*-**g***a*o	→	[ɸuɾ*u*-**ɣ***a*o]
c.	t→d	/ʃabo*ɴ* + **t**ama/	→	ʃabo*ɴ*-**d**ama	N/A	[ʃabo*ɴ*-**d**ama]
d.	k→g	/ʃi*ɴ* + **k**ao/	→	ʃi*ɴ*-**g**ao	N/A	[ʃi*ɴ*-**g**ao]

For words with underlying voicing, if the underlying manner of articulation in the Japanese loanword is assumed to (approximately) match that of the English source words, English [d, g] would be borrowed into Japanese as /d, g/. Thus, English *dance* [**d**æns] and *guard* [**g**ɑɹd] would correspond to Japanese /**d**ansu/ and /**g**aado/. The fact that these words appear as [ɸook*u*-**ð***a*ɴsu] 'folk dance' and [bod*ii*-**ɣ***aa*do] 'body guard' when compounded suggests the same kind of intervocalic lenition process must be operative. In other positions (word-initial and post-nasal), lenition does not apply, hence the realization as stops [d, g] is maintained. The derivation of the lenited words is therefore as in (29a-b), as opposed to the non-lenited words in (29c-d):

(29)

			Lenition	
a.	d→ð	/ɸook*u*-**d***a*ɴsu/	→	[ɸook*u*-**ð***a*ɴsu]
b.	g→ɣ	/bod*ii*-**g***aa*do/	→	[bod*ii*-**ɣ***aa*do]
c.	d→d	/ɾate*ɴ*-**d**aɴsu/	N/A	[ɾate*ɴ*-**d**aɴsu]
d.	g→g	/suki*ɴ*-**g**aado/	N/A	[suki*ɴ*-**g**aado]

To summarize, if the underlying forms of the words in (28) and (29) are taken into consideration, it becomes clear that intervocalic lenition is responsible for the alternations documented in Experiment 2. The fact that the same pattern holds for *Rendaku* words and loanwords alike suggests that this lenition process is independent of *Rendaku* itself. At a higher level, these results mean that the instances of *dʒ→ʒ* and *b→β* lenition observed for loanwords in Experiment 1 are not an isolated phenomenon but rather reflect a more general process of intervocalic lenition that includes *d→ð* and *g→ɣ* as well. That is, it applies to *all* voiced non-continuants in Japanese.

2.3. General discussion:

When the results from the two experiments are taken together as a whole, four major findings emerge. First, the present study successfully confirms how continuancy is modulated by position, which was reported in the literature only sporadically and partially, or sometimes downplayed, as summarized in Sections 1.2.2-1.2.3. More specifically, data from both types of measurements (visible bursts and frication duration) across both experiments led to the conclusion that intervocalic position is associated with a longer frication duration and a lower rate of bursts while post-nasal position is associated with a shorter frication duration and a higher rate of bursts — cf. (11) and (13) for frication duration and (14)-(15) and (23)-(24) for bursts. This suggests that voiced obstruents in present-day Japanese are predominantly produced as the lenis/continuant sounds *z*, *ʒ*, *β*, *ð*, *ɣ* in intervocalic position and as the fortis/non-continuant sounds *dz*, *dʒ*, *b*, *d*, *g* in post-nasal (and word-initial) positions.[18] Notably, this positional contrast is observed irrespective of the obstruents' place of articulation and their source of voicing, i.e., whether it is underlying or due to *Rendaku*.

Second, the theoretical discussion in Sections 2.1.5 and 2.2.5 made it clear that, when the underlying forms are taken into consideration, lenition and fortition both need to be postulated, as depicted in (20)-(21) and (28)-(29). Notably, the combined results of Experiment 1 and Experiment 2 have led to the conclusion that intervocalic lenition in Japanese is quite general and applies to *all* of the underlying voiced non-continuants (*dz*, *dʒ*, *b*, *d*, *g*). Treating post-nasal fortition likewise as affecting all of the underlying voiced continuants (*β*, *z*, *ʒ*) naturally leads to the proposed *Rendaku* application in *ha-gyoo* sounds, i.e. /*ɸ*/→*β*(→*b*). Under this analysis, fortition and lenition are independently motivated processes (separate from *Rendaku*), hence *Rendaku* can be maintained as a straightforward voicing rule. Thus, the present analysis reproduces one of the key advantages of the traditional analysis discussed in Section 1.1.

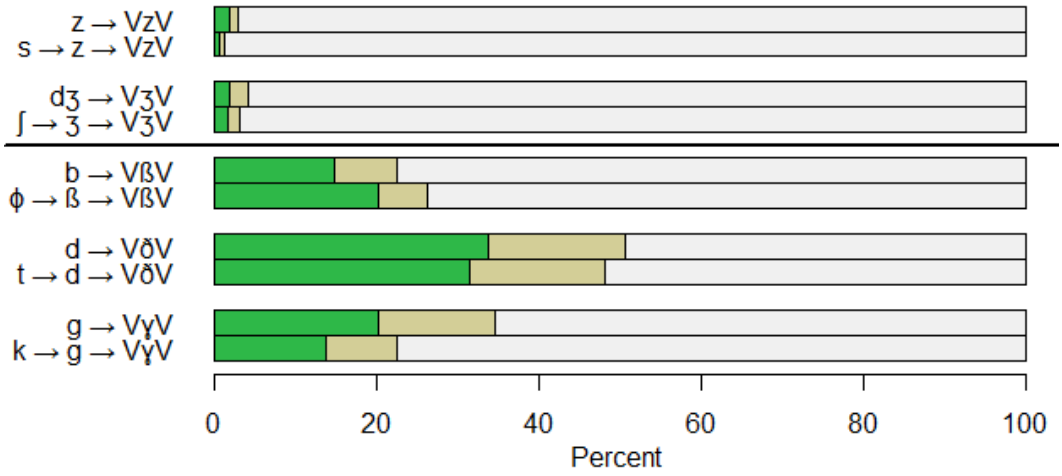
Third, the observed positional contrasts are not categorical in nature. Note, for example, the overlap in the distributions of frication duration in (11) and (13) between the two positions. Similarly, note that, in (14)-(15) and (23)-(24), the contrast in visible bursts was not always, say, 0-10% bursts intervocalically and 90-100% otherwise. Instead, for example, among tokens of *b*/*d*/*g*, bursts were visible intervocalically 15-45% of the time and post-nasally 30-60% (as well as word-initially 70-75%). These facts mean that there is some extent of variation even within a single position such that voiced obstruents are sometimes pronounced as continuants post-nasally and as non-continuants intervocalically (contrary to the general trend). Thus, the contrast in question appears to be more gradient in nature than what would be expected from a simple phonological rule of positional allophony, unconditionally changing one sound into another. This finding reinforces previous reports of the variability of these alternations that have been made in the literature (discussed in Sections 1.2.2-1.2.3 above).

Fourth and finally, the two subclasses of obstruents divided by the thick horizontal lines in (30) and (31) below pattern differently. (These figures are identical to (14)-(15)

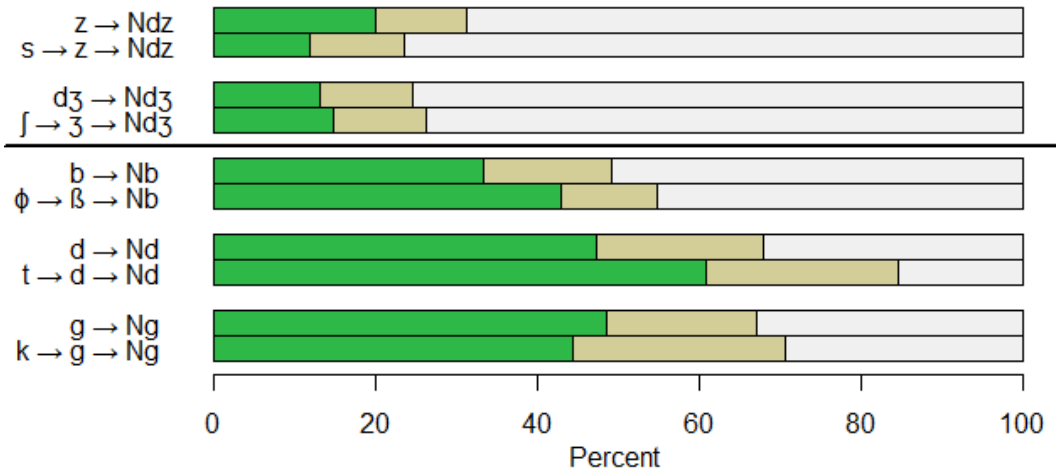
[18] Since word-initial position was not included in Experiment 1, which was designed to examine only the output of *Rendaku* voicing, data is unavailable for *b* (or *z* or *ʒ*) in this position. It may be expected, for example, that the rates of bursts for *b* in this context would pattern similarly to that of *d* and *g*, but this remains to be verified in future research.

and (23)-(24) except, first, that phonological changes are explicitly indicated with arrows, and second, that the bars have been regrouped and divided by the line to indicate the two subclasses in question.)

(30) Combined results for visible bursts (*z*/*dz*, ʒ/*dʒ* and *β*/*b*, *ð*/*d*, *ɣ*/*g*): Intervocalic position



(31) Combined results for visible bursts (*z*/*dz*, ʒ/*dʒ* and *β*/*b*, *ð*/*d*, *ɣ*/*g*): Post-nasal position



Specifically, *z*/*dz* and ʒ/*dʒ* systematically showed fewer bursts than *b*/*β* , *d*/*ð*, and *g*/*ɣ*. This suggests that the former is realized as lenis/continuant more frequently than the latter. This asymmetry held true of both post-nasal and intervocalic positions (and therefore runs orthogonal to the positional contrast just discussed).

By uncovering these four findings, the present study's two experiments provide a clearer picture of the nature of Japanese fortis-lenis manner alternations. What remains unclear, however, is *why* and *how* each of these four facts has arisen. More specifically, it is an open question: (i) whether there is independent motivation to postulate these positional fortis-lenis alternation processes in the phonology of present-day Japanese, (ii) whether such processes are indeed general enough to affect *all* the voiced obstruents indiscriminately, (iii) why the effects of these alternations are continuous and noncategorical in nature, and (iv) why *z*, ʒ, *dz* and *dʒ* show more lenis pronunciations than the other voiced obstruents.

Note that the above four findings (and the four new questions just mentioned) arose

out of experimental data from the speakers of present-day Japanese. In the pursuit of explanations as to why these synchronic facts hold true, the approach pursued in this work (namely Sections 3 and 4 below) is to look to diachrony for answers. In other words, it will be examined how these four findings 'make sense' diachronically — and hence how the four new questions can be answered. Ultimately, it will be pointed out that a wealth of historical evidence lends support to the proposed analysis of *ha-gyoo* sounds ($\phi \rightarrow \beta \rightarrow b$).

3. Overview of historical changes:

Describing the synchronic grammar of a language involves "slicing" the continuously changing grammar at one particular point in time and examining the cumulative effects of diachronic changes up to that point (cf. Hamada 1960:20). From this perspective, in order to understand exactly how the output of *Rendaku* came to be phonetically realized and transcribed with *ha-gyoo* sounds in present-day Japanese, it is essential to investigate the evolution of the language's sound system and the writing system. In each of Sections 3.1-3.3 below, the philological literature will be briefly reviewed in an attempt to sort out the varied (and often conflicting) views reported on three issues of diachronic changes in the Japanese language, namely: (i) how voicing, including that induced by *Rendaku*, has manifested itself throughout history, (ii) when and how the *ha-gyoo* consonants changed, especially in relation to *p*, and (iii) when fortition and lenition were introduced into the grammar and how they have maintained their effects since then with gradual changes in the ways they apply. Then, starting in Section 4.1, a bird's-eye-view analysis and discussion of these diachronic investigations will be provided, unraveling how these phonological events interacted with the orthography and came to be represented (or not represented) in writing. Ultimately, it will become clear how the three diachronic changes just listed are reflected in the phonetic realization of the *ha-*/*ba-*/*pa-*gyoo sounds in present-day Japanese in precisely the way expected from the proposed analyses of *Rendaku* voicing, and in precisely the way confirmed by the experimental results from Section 2. While the historical investigations in Section 3 provide a somewhat long deviation from the discussions in the previous sections, they will turn out to provide the key to the satisfactory understanding of the phonetic phenomena in present-day Japanese when they are examined in the light of the interactions between diachrony and synchrony of the Japanese language in Section 4.

3.1. Diachrony of voicing:

It is well known that, from ancient to medieval Japanese, voicing in obstruents was quite restricted both in distribution and function. Until around the 8th century (8C), in native Japanese words, voiced obstruents generally did not occur in initial position, suggesting voicing was not contrastive in word-initial position (Tsukishima 1964:32, Nakata 1972c:24, Komatsu 1981:60, 104). Given this phonotactic restriction against initial voiced obstruents, it is widely believed that *Rendaku* marked compoundhood in word-internal position, as illustrated in (32) (Yamada 1904a, Komatsu 1981:104, Kubozono 1999:117).

(32) a. **k**iɾi 'mist'
b. ***g**iɾi 'mist'
c. asa-**g**iɾi 'morning mist'

Rendaku has been operative throughout the history of Japanese since at least the *Nara* period (8C) (Nakata 1972b:28, Takayama 2015) up through to present-day Japanese. This turns out to be a crucial historical fact that demands full attention, as will be discussed in Section 4.1 below.[19]

The overwhelming majority of the modern philological literature suggests that voicing of obstruents in Japanese was not rigidly contrastive, at least in everyday language (Hamada 1977:590, Takayama 1992-3:120), until somewhere between the Heian period (ca. 9~12C) (Mabuchi 1977:272-3) and the early Edo period (ca. 17C) (Hamada 1960:24). This view is often justified with the following two observations (Okumura 1972:112, Numoto 1992:78). First, voicing in obstruents was not indicated on *kana* in writing – for example, even in *iroha-uta* 'Japanese alphabet song' (10~11C). Second, independent *kana*s were never invented for transcribing voiced obstruents.

Mabuchi (1977:272-3) argues that before voicing was contrastive, it created only allophonic variation, providing *mu**dz**ukaʃii* 'difficult' and its alternant *mu**ts**ukaʃii* in modern Japanese as an example remnant of this allophonicity. He and Hamada (1977:590-1) independently argue that voicing first became contrastive in Chinese loanwords, perhaps initially among Buddhists who were familiar with a similar distinction in Sanskrit (Mabuchi 1977:271-3). According to these authors, the contrastiveness of voicing within this limited set of vocabulary did not immediately have a broad impact on the language as a whole, i.e., it was only later that voicing contrast gradually spread to the general public and to everyday vocabulary. As such, the allophonic nature of voicing remained unaffected for quite a while.

Okumura (1972:112) argues that this state of affairs continued well beyond the Heian period (9~12C), citing the observation that the voicing distinction was rarely transcribed even in the documents from the post-Heian period. Similarly, Kiyose (1985:84-85) points out that voicing on obstruents was neglected even during the Muromachi period (14~16C), citing riddles from *Gonarain-gyosen-nazo* (1516) that make sense only when the voicing on certain obstruents is disregarded. Thus, the philological literature suggests that voicing had only a weak contrastive function at any position in native Japanese words possibly until late 16C or as late as 17C, at which point it became more clearly contrastive. This timing of when voicing became contrastive turns out to be important in the discussion below where orthography is brought into the picture.[20]

---

[19] Many researchers (e.g., Hizume 2003) consider that *Rendaku* historically stemmed from prenasalization of obstruents in ancient Japanese, e.g., *asa-**ⁿk**iri*, which itself is said to have arisen from the genitive marker *-no*. Such prenasalization, as well as the resulting spread of voicing, is often observed in present-day *Toohoku* dialects (Kiyose 1985:79), e.g., *ma**ⁿd**o* < *ma**t**o* 'target' (cf. Miyashita, et al. 2016). Taking this historical change into consideration, McCawley (1968:86-87) proposes that *Rendaku* consists of a set of rules that turn obstruents first into "prenasalized obstruents" and then into various voiced sounds. Some also regard *Rendaku* as a phonological process involving voicing assimilation between two voiced sounds (Yamada 1904b, Kubozono 1999:43, 114). The present paper is not concerned with the properties of *Rendaku* beyond voicing, nor with various restrictions known to be imposed on its application. See Ito and Mester (1986), Kubozono (1999), Rosen (2001), Vance and Irwin (2016), among others, for relevant discussion.

[20] Two apparent counterexamples are worth noting. First, although voicing is distinguished in *Manyoo-gana* (Chinese characters used to represent the sounds of Japanese, e.g., 久 **k**u vs. 具 **g**u) in documents from the *Nara* period (8C), many researchers assume they were recorded by and for scholars from China (Hamada 1977:590, Kiyose 1985:78) and possibly also Korea (Numoto 1992:81). Numoto (1992:78) also

3.2. Diachrony of *ha-gyoo* consonants

3.2.1. The rise of *ɸ* and its subsequent changes to *h* and *ç*:

There appears to be a consensus among philologists that all *ha-gyoo* consonants were pronounced as *p* in native words in ancient Japanese. For this view, see Hoffmann (1868), Ueda (1903:39), Tsukishima (1964:28), Nakata (1972b:25), Toyama (1972:194), Komatsu (1981:249), among others. It is also well attested that, later in history, *p* came to be lenited and replaced by *ɸ* (except in onomatopoeia). There is no consensus, however, as to when this diachronic change took place. Hashimoto (1928:201-2), for instance, argues that the *p*-to-*ɸ* change took place by the *Nara* period (8C) — a view also endorsed by Tsukishima (1964:28), Nakata (1972b:25) and Sato (1977:237), among others. Others like Kamei (1960:128-129) and Kiyose (1985), however, maintain that *p* remained much longer. Kiyose (p. 84), for example, argues that the change from *p* to *ɸ* took place sometime during the *Muromachi* period (14~16C), at least in the western part of Japan. In either case, there exist documents suggesting that the consonant in question had come to be pronounced as *ɸ* at least by the mid to late *Muromachi* period, i.e., by around 15C. For example, in works written by Portuguese Christian missionaries, such as and *Vocabviario da Lingoa de Iapam com Adeclaração em Portugues* (日葡辞書: 1603) and João Rodriguez's *Arte da Lingoa de Iapam* (日本大文典: 1604~8), the *ha-gyoo* sounds were transcribed with *f* irrespective of the following vowel. It is considered by Toyama (1972:194-5) and Hashimoto (1966:70-71, 109) that the bilabia1 fricative *ɸ* was transcribed approximately with <f>, i.e. the glyph normally reserved for the labio-dental fricative in late 16th-century Portuguese.

(33)	Christian records	Gloss	Present-day Japanese
a.	**F**ato	'pigeon'	**h**ato
b.	**F**ito	'person'	**ç**ito
c.	**F**uyu	'winter'	**ɸ**uyu
d.	**F**ebi	'snake'	**h**ebi
e.	**F**oka	'other'	**h**oka

Presumably, the historical change in question gradually spread from one region to another, thereby causing geographical variation as well as time lag. More specifically, the *p*-to-*ɸ* change may have started as early as 8C on a small scale in the west and gradually spread toward the east and became more pervasive by 15~16C. In fact, Kiyose (1985:86) points out that *ha-gyoo* consonants were transcribed with *Hangul* in *Choosen-ban Iroha* (The Korean Edition of *Iroha*: 1492) sometimes as plosives and other times as fricatives, and argues that this may indicate dialectal and/or free variation between *ɸ* and *p* in 15C.[21]

---

points out that the documents written in *Manyoo-gana* were basically only for official use and did not reflect the language used by the general public. Second, while word-initial voiced obstruents in Japanese started to appear in records from the *Heian* period (9~12C), Okumura (1972:111) argues that most of them were derived phonologically due to the deletion of word-initial high vowels, e.g., ***d****aku* < ***i****daku* 'hold' (both of which remain in the present-day vocabulary). As such, the word-initial phonotactic restriction on voiced obstruents may still have been in effect at the time.

[21] Many researchers consider that, after replacing *p*, *ɸ* underwent further phonological change *ɸ*→*w* (called *ha-gyoo tenko*) word-internally, which started toward the end of 8C and came to be widely spread by 11C (Tsukishima 1964:29, Sato 1977:248). Hayata (1977:359) considers that *ɸ* lenited from *p* was voiced and

Although *p* came to be replaced by *ɸ* in most of the Japanese vocabulary, it tended to remain in onomatopoeia. Given the imitative function of mimetics, their resistance to sound change is understandable. Many such instances of *p* in onomatopoeia have been inherited even in present-day Japanese (Toyama 1972:197, Komatsu 1981:269). For example, Kubozono (1999:61-62) cites the cognate pairs in (34) below, whose pronunciation differences and etymological connections make sense given the co-existence of the *p*-to-*ɸ* change and the resistance to this change in onomatopoeia.

(34) a. **p**ikaɾi 'flash' — **ç**ikaɾi 'light'
b. **p**iyo-**p**iyo 'peep' — **ç**iyoko 'chick'

It should also be pointed out, however, that while onomatopoeia tended to resist the *p*-to-*ɸ* change, it is not the case that they *never* underwent this change. This fact is evidenced by words such as those in (35) below, for which both pronunciations (with *p* as well as *ɸ*) have been attested in writing since 15C-16C, suggesting that they coexisted during and after the *p*-to-*ɸ* change took place.[22]

					Written attestations[23]
(35)	a.	(i)	***ɸ***ika-***ɸ***ika	'flash'	1534,[24] 1595, 1632, 1717, 1768, 1809~13
		(ii)	**p**ika-**p**ika	'flash'	1841~2, 1887, 1903, 1917, 1931, 1969, 1970
	b.	(i)	***ɸ***iyo-***ɸ***iyo	'peep'	Late 10C, Mid-Late 16C, 1792, 1887, 1932
		(ii)	**p**iyo-**p**iyo	'peep'	1688~1710, 1887, 1903

Since the replacement of *p* by *ɸ* was so pervasive in its position in the major lexical strata, it is perhaps best regarded as the wholesale replacement of an entire phoneme. This diachronic change may be regarded as a case of lexical diffusion, whereby lenition spread within non-onomatopoetic vocabulary and eventually marginalized *p* in the phonemic inventory of native and Sino-Japanese words.[25]

The *ɸ* sound then gradually underwent further changes, most notably during the first half of the *Edo* period (17~19C). Tsukishima (1964:34), Nakata (1972a:34), Okumura (1972:126-8), Toyama (1972:244), and Komatsu (1981:249-253) all argue that *ɸ* came to

---

became *β* word-internally by 11C, which easily turned into *w* with slight laxing of frication. See also Frellesvig (2010:203) for an alternative analyses *p*→*b* and *β*→*w* without involving the intermediate stage *ɸ*. Because of *ha-gyoo tenko* (and other additional change), *ɸ* and its allophonic variants generally do not appear at the non-initial position of a morphologically simplex word though exceptions are known to exist (e.g., *ha**ha*** 'mother'). See Hizume (2002:98) for more examples and a relevant discussion.

[22] The written attestation dates for the *p*-versions in (35a-b) are so late (17C at the earliest) presumably because the *p* sound was not represented orthographically (using the *handakuten* diacritic) until around that time, as will be discussed in Section 4.6 below. The earliest attestation date of "Late 10C" for (35b-i), on the other hand, may suggest that the *p*-to-*ɸ* change had already started to take place, on whatever scale, by that point in time.

[23] Unless otherwise indicated, all of the attestation dates provided in the present paper are cited from the 2nd edition of *Nihon Kokugo Daijiten* (2007), one of the most comprehensive resources for historical data on the Japanese language. Attestation dates indicated here are limited to citations where the pronunciation in question, e.g. *ɸ* vs. *p*, can be confirmed from orthography (either from a *kana* or Roman letter). Allophonic variations within the *ha-gyoo* sounds (e.g. *h* vs. ç) are not indicated.

[24] The attestation date 1534 from *Shiganikkai* (四河入海) listed in (35 a-i) accords with its generally accepted compilation date rather than what *Nihon Kokugo Daijiten* indicates (Early 7C).

[25] We would like to thank Stuart Davis for suggesting this possibility.

be pronounced as *h* in most dialects during 17C. Then, starting in the eastern dialects, *ɸ* came to be pronounced as *ç* before *i* in early 18C. These changes, recapitulated in (36) below, led to the varied pronunciation of the *ha-gyoo* sounds in present-day Japanese.[26]

(36)		During 17C	Early 18C
a.	**ɸ**a	→ **h**a	
b.	**ɸ**i	(→ **h**i)	→ **ç**i
c.	**ɸ**u		
d.	**ɸ**e	→ **h**e	
e.	**ɸ**o	→ **h**o	

Debuccalization may have applied to *ɸ* for reasons of economy-of-effort — namely, by eliminating all supralaryngeal articulation. Sequences of *ɸu* as in (36c) may have been the sole exception since *ɸ* and *u* both involve the lips as an articulator (for lip compression in the case of *u*) and, as such, these sequences may require relatively less articulatory effort compared to sequences with the other vowels. Then, in the early 18C, palatalization applied to *hi* (or *ɸi*) sequences and altered the *h* (or *ɸ*) to *ç*, apparently for the assimilation of the tongue position.[27] Thus, the set of changes described in (36) constitute a valid explanation as to the diachronic source of the diverse pronunciation of the *ha-gyoo* sounds (*ha*, *çi*, *ɸ*u, *he*, *ho*) in present-day Japanese, as exemplified by the words first presented in (2) above: ***h**ana* 'flower', ***ç**iki* 'pulling', ***ɸ**ɯro* 'bath', ***h**eja* 'room' and ***h**o*ri 'carving'.

Thus, although *ɸ* may initially seem to be phonetically less likely as an underlying sound than *p*, it is perfectly feasible to regard the underlying *ɸ* analysis of the *ha-gyoo* consonants in present-day Japanese as a legacy of a series of diachronic changes just described (/p/→/ɸ/→[h, ç, ɸ]).

3.2.2. Re-introduction of *p*:

While the *ha-gyoo* sounds in Japanese were still undergoing the changes just described above, the history of these sounds took another interesting twist. According to several researchers (e.g., Hamada 1954:24-5, Okumura 1972:90-91, Asada 2000:825), while *p* was being replaced by *ɸ*, it was gradually reintroduced into the non-onomatopoetic vocabulary due to influence from Chinese. Loanwords from Chinese containing *p,* as in (37) below, are generally believed to be the trigger for the reintroduction of *p*.

(37)					Written attestations
	a.	kop**p**oo	(骨法)	'frame'	751, 1571, 1603-4
	b.	eɴ**p**a	(煙波)	'foggy water surface'	827, ca.1220, 1603-4

According to Wang's (2012) interpretation of *Guangyun* (廣韻, an influential Chinese rhyme dictionary completed in 1008), the loanwords in (37) were pronounced in Chinese as [kuət-**p**ɐp] and [ʔen-**p**uɑ], respectively, during or before the Tang dynasty (ca. 7~10C).

---

[26] Because of *ha-gyoo-tenko* (*ɸ* → *w*) mentioned in footnote 21, these phonological changes were generally limited to word-initial position.

[27] There appears to be no consensus on whether *ɸ* changed into *ç* directly or the change was mediated by *h*. For example, Komatsu (1981:252-3) suggests that the latter is likely while Hashimoto 1966:116) maintains the possibility of the former, at least in some dialects.

Although there is no direct evidence, such words are thought to have been pronounced with *p* when they were loaned into Japanese (Okumura 1972:90).[28] Moreover, pronunciations with *p* are also attested in later documents and maintained in present-day Japanese. The *p* sound then gradually spread from Chinese loanwords to Sino-Japanese words, replacing *ɸ* in certain contexts within compounds as in (38).

					First attestation	
(38)	a.	kiɴ**p**eɴ	(近辺)	'neighborhood'	1604~8	cf. ***ɸ***eɴ (辺)
	b.	ɾip**p**uku	(立腹)	'getting angry'	1776	cf. ***ɸ***uku (腹)

The pronunciation with *p* in Chinese loanwords and Sino-Japanese words is confirmed most clearly in the Romanized transcriptions used in various Christian documents published in 16~17C (Okumura 1972:90). It is also pointed out in the philological literature that *p* appeared in instructions for reciting various types of texts aloud from as early as 13C (e.g. Toyama 1972:197-8, Asada 2000:833).

An independent (but presumably connected) development of the *p* sound in native Japanese words occurred in war tales from the *Kamakura* period (13~14C) like *Heeke-monogatari* (1219～1222). In words such as those in (39a-b) below, *ɸ* was sometimes replaced with (geminate) *p* in order to encourage vivid and powerful recitation (Toyama 1972:197, Komatsu 1981:279-282). Similar pronunciations are also attested in commentary books written in colloquial Japanese like *Goongyokujyoojyoo* (1429~41), *Shikishoo* (1477), and *Shuekishoo* (1477), as exemplified by the words in (39c-d).

(39)	a.	yoku **ɸ**iite	→	yop**p**iite	'drawing (a bow) to the full'	(*Heeke-monogatari*)
	b.	a**ɸ**aɾe	→	ap**p**aɾe	'admirable'	(*Heeke-monogatari*)
	c.	mo**ɸ**aɾa	→	mop**p**aɾa	'entirely'	(*Goongyokujyoojyoo*)
	d.	ya**ɸ**aɾi	→	yap**p**aɾi	'still/as expected'	(*Shikishoo*, *Shuuekishoo*)

Presumably, some western words loaned later on in history, such as those from Portuguese in (40) below, further reinforced the reintroduction of *p*.

							First attestation
(40)	a.	*kap**p**a*	'raincoat/gown'	<	*ca**p**a*	'jacket'	ca. 1615
	b.	*koɴ**p**assu*	'compass'	<	*com**p**asso*	'compass'	1709

Most of the words containing *p* in (37)-(40) still remain in the vocabulary of present-day Japanese.

As a consequence of the reintroduction of *p* just discussed, the philological literature often presents the view (without providing concrete arguments) that *p* re-emerged as an independent phoneme after 13C (Okumura 1972:90, Toyama 1972:198, Komatsu 1981:277, Numoto 1990:8, Numoto 1992:87). In Section 4.4 below, the phenomena in (37)-(40) will be re-examined, and an alternative proposal will be sketched out.

3.3. Diachrony of the fortition and lenition of voiced obstruents:

Several changes are known to have taken place in the phonology of medieval to early modern Japanese (ca. 13~17C) that affected the phonetic realization of voiced obstruents. These changes have important implications for interpreting the data presented in Section

[28] In the historical sources for the earliest attested date in (37a) and (37b), the words in question are written in Chinese characters, hence the pronunciation of these words cannot be confirmed.

2 on the phonetic realization of voiced obstruents in present-day Japanese (including the *Rendaku*-voiced *ha-gyoo* consonants). First, many researchers consider that, in medieval Japanese, the dental stop *d* underwent both affrication and a recession in place-of-articulation in front of high vowels, as illustrated in (41).[29]

(41) a. **d**i → **dʒ**i
b. **d**u → **dz**u

These changes reduced the acoustic distance between *di/du* (now realized as *dʒi*/*dzu*) and *ʒi*/*zu*. This historical incident paved the way to a widespread but chaotic merger of these fricatives and their corresponding affricates (i.e., *ʒi* ≈ *dʒi* and *zu* ≈ *dzu*) in medieval Japanese. This merger is reflected in the historical variation attested for words like *su**dʒi*** 'street' pronounced as *su**ʒi*** (スヂ~スジ) and *ku**dzu*** 'trash' pronounced as *ku**zu*** (クヅ~クズ) (Takayama 2015:407). While the pronunciation of these sounds came to be mixed-up, the *kana*s representing them remained the same, thus making the relevant *kana*-pronunciation correspondences opaque and unstable — a phenomenon known as the "confusion of the four *kana*s" (四つ仮名の混同).[30]

This *kana*-pronunciation entanglement was advanced by three additional factors. First, there is historical evidence that lenition and fortition both applied in a phonologically conditioned manner at one point in the history of Japanese. *Kenshukuryookoshuu* (蜆縮凉鼓集:1695), an instructional book promoting prescriptive pronunciation, describes pronunciations with affricates (*dʒ*/*dz*) in words like (42a) below as "correct" in contrast to pronunciations with fricatives (*ʒ*/*z*) in words like (43a). A similar comparison is also made between the "correct" pronunciation with fricatives (*ʒ*/*z*) in words like (42b) and "incorrect" pronunciations with affricates (*dʒ*/*dz*) in words like (43b) (Kamei 1950:82, 85, Toyama 1972:249, Okumura 1972:102). These examples are structured such that the second morpheme is identical between the pairs of bimorphemic words in (42) and (43), as indicated by the Chinese characters, e.g., <着> appearing in the pair (42a-i) and (43a-i).

[29] See Takayama (2015:3.1) for further discussion of the phonological factors behind the allophonic variation that arose in all of *ta-/da-gyoo* sounds and *sa-/za-gyoo* sounds. Mainly based upon the documents written by Chinese and Korean scholars in 16C, many scholars have concluded that the affrication in (41) took place in 16C (Okumura 1972:124, 126, Toyama 1972:190, Takayama 2015:406). See footnotes 34 and 39 below for a brief discussion on this conclusion.

[30] The same problem also arose in the corresponding *yooon* (phonologically palatalized sounds), namely ***ʒa*** vs. ***dʒa*** (ジャ~ヂャ), ***ʒu*** vs. ***dʒu*** (ジュ~ヂュ), and ***ʒo*** vs. ***dʒo*** (ジョ~ヂョ). See also Sato (1977:250), Takayama (2003) and Takayama (2015:3.3) and the references cited therein, *inter alia*, for discussion on how the confusion of the four *kana* characters may be related to the prenasalization of voiced obstruents (a topic which will not be addressed in the present work).

(42) **"Correct" pronunciation**:

a. **Affricates**:	(i)	還**着**	ge**ɴdʒ**aku	'return to work'	cf. **tʃ**aku	(着)
	(ii)	平**地**	hei**dʒ**i	'flat ground'	cf. **tʃ**i	(地)
	(iii)	弓**杖**	juɴ**dz**ue	'bow as a cane'	cf. **ts**ue	(杖)
	(iv)	神**通**	ʒiɴ**dz**uu	'supernatural power'	cf. **ts**uu	(通)
b. **Fricatives**:	(i)	行**者**	gjau**ʒ**a	'ascetic'	cf. **ʃ**a	(者)
	(ii)	藤**氏**	tou**ʒ**i	'the Fujiwara clan'	cf. **ʃ**i	(氏)
	(iii)	奇**瑞**	ki**z**ui	'auspicious omen'	cf. **s**ui	(**瑞**)
	(iv)	香**水**	kau**z**ui	'fragrant water offering'	cf. **s**ui	(水)

(43) **"Incorrect" pronunciation**:

a. **Lenition**:	(i)	執**着**	ʃiɸu**ʒ**aku	'adherence'	cf. **tʃ**aku	(着)
	(ii)	空**地**	kuu**ʒ**i	'open space'	cf. **tʃ**i	(地)
	(iii)	竹**杖**	take**z**ue	'bamboo cane'	cf. **ts**ue	(杖)
	(iv)	普**通**	ɸu**z**uu	'generality'	cf. **ts**uu	(通)
b. **Fortition**:	(i)	判**者**	haɴ**dʒ**a	'judge'	cf. **ʃ**a	(者)
	(ii)	源**氏**	geɴ**dʒ**i	'the Minamoto clan'	cf. **ʃ**i	(氏)
	(iii)	天**瑞**	teɴ**dz**ui	'auspicious omen'	cf. **s**ui	(**瑞**)
	(iv)	神**水**	dʒiɴ**dz**ui	'water offering'	cf. **s**ui	(水)

The fact that the pronunciations of the highlighted obstruents in (43) were regarded by some intellectuals as "incorrect" may imply that one or more new phonological processes had begun to take hold and alter the obstruents' phonetic realization (in contrast to the more conservative variety in (42) lacking these newly-introduced phonological processes).[31] More specifically, lenition had begun to apply in intervocalic position to the underlying affricates in (43a), and fortition had begun to apply in post-nasal position to the underlying fricatives in (43b). Thus, these data suggest that there existed a period during which lenition and fortition started and continued to apply systematically in a position-modulated manner within certain varieties of Japanese, thus obscuring *kana*-pronunciation correspondences even further.[32]

Second, historical records show that these lenition and fortition processes *themselves* became unstable over time. More specifically, the philological literature suggests that the *dʒi-ʒi* and *dzu-zu* distinctions had both come to be obscured (at least in certain words)

---

[31] In most of the examples from *Kenshukuryookoshuu* reported in (42a), the affricates (*dʒ*/*dz*) appear in a post-nasal position. The same affricates, however, did appear in an intervocalic position as well, as shown by the examples from the same era in (i) (along with their written attestation dates).

(i)	a.	愛**着**	ai**dʒ**aku	'attachment'	(1603~4)
	b.	意**地**	i**dʒ**i	'will/ego'	(1563, 1603~4, 1641)
	c.	**側杖**	soba**dz**ue	'embroilment'	(1698, 1783)
	d.	弘**通**	gu**dz**uu	'spread of Buddhism'	(1591, 1603~4, 1705)

[32] The author of *Kenshukuryookoshuu* sets out to describe the "constantly exercised incorrect pronunciations" (i.e., pronunciation mistakes) made by the people of Kyoto (in western Japan, the capital of the country at the time), suggesting that this may have been a geographically limited phenomenon.

nearly everywhere in Japan sometime between 16C and 17C (e.g., Tsukishima 1964:33, Hashimoto 1966:102-7, Okumura 1972:102−3, Toyama 1972:198-202, 247-251, Komatsu 1981:127, Takayama 2015:3.2). The following examples from *Arte da Lingoa de Iapam* (日本大文典:1604−8) are among those typically cited, illustrating unexpected lenition in (44) and unexpected fortition in (45). Note the mixture of the contexts for the same phonological change (V_V and #_ in (44)) and (N_, #_ and V_V in (45)). The author (João Rodriguez) contrasts the expected pronunciations in the Kyoto dialect (to the left of the arrows) with the "confused pronunciations" thereof (to the right of the arrows) and describes such pronunciations as a defect of the otherwise respectable Kyoto dialect. (The original Romanized transcriptions have been converted to phonetic symbols here.)[33]

(44) *dʒi*/*dzu* (ヂ/ヅ) → *ʒi*/*zu* (ジ/ズ):

		Expected		Confused	
a. 水	'water'	*Mi**dz**u*	→	*Mizu*	(V__V)
b. 此の中	'lately'	*Cono **dʒ**iŭ*	→	*Cono **ʒ**ŭ*	(V__V)
c. 地盤	'ground'	***dʒ**iban*	→	***ʒ**iban*	(#__)
d. 直に	'soon'	***dʒ**iki-ni*	→	***ʒ**iki-ni*	(#__)

(45) *ʒi*/*zu* (ジ/ズ) → *dʒi*/*dzu* (ヂ/ヅ):

		Expected		Confused	
a. 本寺	'head temple'	*Fon**ʒ**i*	→	*Fon**dʒ**i*	(N__)
b. 自然	'automaticity'	***ʒ**ine*ɴ	→	***dʒ**ine*ɴ	(#__)
c. 参らず	'not come'	*Mai**rz**u*	→	*Mai**rdz**u*	(V__V)

It is generally believed that this 'blurring' of the *dʒi*-*ʒi* and *dzu*-*zu* distinctions first started in the eastern dialects and gradually spread to the west, excluding some dialects in Kyuushuu and Shikoku. *Wajishooranshoo* (1693), for example, reports a contrast between the dominance of lenition in the Kyoto dialect (in western Japan) and the dominance of fortition elsewhere (Toyama 1972:248-250).[34]

Finally, in the modern era, *kana*-pronunciation correspondences were obscured even further due to the Japanese government's adoption of policies on how the four *kana*s in question should be used (in an attempt to mitigate the chaos). At first, the *rekishiteki kanazukai* (literally, 'historical *kana* usage') system adopted during the *Meiji* period appealed to the etymology-based ancient pronunciation of each word. This recommendation was later overturned, however, in the *Gendai Kanazukai* (現代仮名づ

[33] These examples are adapted from Doi's (1955:608) Japanese translation of *Arte da Lingoa de Iapam*. The following is the correspondence between the romanization in Christian records and their pronunciations generally assumed in the literature: *g* = [dʒ], *i* = [ʒ], *dz* or *zz* = [dz], *z* = [z]. *Cono **j**ŭ* in (44c) and *Fon**j**i* in (45a) were not written *Cono **i**iŭ* (**i** = [ʒ]) and *Fon**i**i* (**i** = [ʒ]), respectively, presumably to avoid confusion from having a sequence of two <i> in a row.

[34] In the eastern dialects, examples of confusion can be found in documents from as early as the mid *Kamakura* period (i.e. ca. 13C) (Tsukishima 1964:33, Okumura 1972:102-103, Toyama 1972:202, 250). If the "confusion of the four *kana*s" started this early, the affrication of *t*/*d* in (41) might have started sometime before 13C (at least outside of Kyoto). This calls into question the common assumption that (41) took place in 16C and laid the foundations for (and hence must have predated) the "confusion of the four *kana*s" (cf. footnote 29).

かいい) 'modern *kana* usage' system adopted after World War II (originally in 1946 and then updated in 1986). In the new system, the *kana*s representing the fricatives ***ʒi*** (ジ) and ***zu*** (ズ) were extended even to the transcription of the affricates ***dʒi*** and ***dzu***, thereby completely replacing the corresponding *kana*s (ヂ and ヅ, respectively) except for in a handful of conventionally established cases.[35] This recommendation most likely reflected the fact that ***ʒi*** and ***zu*** were already becoming increasingly frequent in words where ***dʒi*** and ***dzu*** were expected. With the enforcement of this orthography reform, then, the *dominance* of the *za-gyoo* in pronunciation was accelerated even further. In this way, the Japanese government's policy added another twist to the *kana*-pronunciation correspondences, leaving the situation somewhat (or perhaps even more) obscure and unstable in present-day Japanese. The consequences of *Gendai Kanazukai* will be discussed further in Section 4.3.

To sum up, the review of the philological literature leads us to learn, first, that new phonological processes of intervocalic lenition as well as post-nasal fortition started to apply and gradually spread among the speakers by 16C, and second, that the positional conditions of these phonological processes began to loosen before 17C. Third and finally, the orthography reform by the Japanese government after World War II accelerated the dominance of lenis consonants in the alternations between *ʒ*/*z* and *dʒ*/*dz*.

4. Proposed synthesis of synchronic and diachronic findings:

This section presents a broader bird's-eye-view perspective, tying together the various diachronic facts just outlined in Section 3 with the experimental findings from Section 2, in order to make the main arguments of the paper.The ultimate goal is to predict the way *Rendaku*-voiced *ha-gyoo* obstruents are phonetically realized in the grammar of present-day Japanese, which itself reflects the interaction and cumulative result of those diachronic changes.

4.1. Extensiveness of the fortis-lenis alternation:

As mentioned in Sections 3.1-3.2, all *ha-gyoo* consonants in Japanese are believed to have been pronounced as *p* in native words from ancient Japanese but came to be pronounced as *ɸ* some time between 8C (*Nara* period) and 15C (mid-*Muromachi* period). In addition, it is also believed that *Rendaku* has continued to apply throughout the history of Japanese since at least the Nara period. It therefore is reasonable to consider that *Rendaku* applied to *p* and derived *b* before the *p*→*ɸ* change took place. It in fact must also be the case that *Rendaku* has applied to *ɸ* and derived *β* ever since *ɸ* replaced *p* (in the non-onomatopoetic vocabulary).

Based on his impression that the *Rendaku*-output of *p* is always pronounced as *b* and never as *β* in present-day Japanese, Komatsu (1981:254-7) has argued that the *Rendaku*-output of *p* has remained to be *b* throughout history and that *Rendaku* never applied to *ɸ* — even after *p* was replaced by *ɸ*. The results of the first experiment in Section 2.1, however, show that Komatsu's premise (that the *Rendaku*-output of *p* is never pronounced as *β*) is unwarranted, thus invalidating his conclusion. On the contrary, it

[35] Since the orthography reform, the affricates *dʒ* and *dz* have been written with *za-gyoo kana*s in loanwords as well, i.e., using ジ/ズ rather than ヂ/ヅ (for *dʒi*/*dzu*, respectively). The loanwords used in Experiment 1 in Section 2.1 (on the *z-dz* and *ʒ-dʒ* alternations) were spelled following this convention.

would be quite unnatural to assume that, after the *p*→*ɸ* change took place and *ha-gyoo* consonants came to be pronounced as *ɸ* (around 15C), *Rendaku* never applied to *ɸ* for some reason. Such an exception would be especially unexpected in light of the fact that *Rendaku* is still productive even in present-day Japanese, as illustrated by the recently coined words (46a-d) below, suggesting that *Rendaku* has continued to apply throughout history without interruption up through to the present day.

(46) *Rendaku* application in newly coined words in present-day Japanese:

a.	**k**utʃi	'mouth/talk'	→	tame-**g**utʃi	(タメ口)	'peer language' (1960~, *Digital Daijisen*)
b.	**k**e	'hair'	→	ʀoɴ-**g**e	(ロン毛)	'long hair' (1990s~, *Nihongo Zokugojiten*)
c.	**s**eme	'attack'	→	gaɴ-**dz**eme	(ガン攻め)	'aggressive attack (in a PC game)' (2000s~, *online sources*)
d.	**s**eme	'attack'	→	goɴ-**dz**eme	(ゴン攻め)	'aggressive skateboarding performance' (2021~, *online sources*)

Notice, for example, the application of *Rendaku* to ***k****utʃi* 'mouth' as in (46a). In a similar vein, speakers of present-day Japanese can apply *Rendaku* continually and naturally in nonce formation (cf. Ito and Mester 2003:149), for example, as in the nonce words (47a-c) (created by the first author and accepted by two other Tokyo speakers).

(47) a. maunto-**g**utʃi 'mounting language'
b. modʒa-**g**e 'shaggy hair'
c. otome-**g**ana 'girly kana'

Thus, the proposed scenario — that *Rendaku* has applied to *ɸ* and derived *β* ever since *p* was replaced by *ɸ* (including in present-day Japanese) — cannot be excluded as a possibility and deserves due consideration. There appears to be no systematic reason or historical evidence prohibiting the exploration of this possibility.[36] If *Rendaku* applies to *ɸ* and derives *β*, then how does the output *β* come to be pronounced as *b* in some contexts? It is proposed here that: (i) the *β*→*b* change was initially triggered when voiced fricatives began to undergo fortition in medieval Japanese as in (43b), (ii) this same fortition process applied indiscriminately to all voiced continuants (*z*, *ʒ*, *β*), and (iii) this same fortition process *still applies* to all voiced continuants in present-day Japanese. This view is compatible with the analysis reached in Section 2.1 based on the results of Experiment 1, repeated here as (48).

[36] This of course does not rule out the possibility that the *Rendaku*-voiced *b* may have remained in some native words that date back to ancient Japanese, possibly, e.g. *koto**ba*** 'word' (written 言羽 in *Manyooshuu*:8C). Even if this is the case, however, *Rendaku*-voiced *b* would most likely also have been affected by the intervocalic lenition discussed soon below.

(48)

	(i) Input: voiceless continuants	(ii) Output: post-nasal Rendaku-voiced non-continuants	(iii) Output: intervocalic Rendaku-voiced continuants
a.	sakura	kaɴ - **dz**akura	yae - **z**akura
b.	**ʃ**ake	giɴ - **dʒ**ake	**ʃ**io - **ʒ**ake
c.	**h**are	nihoɴ - **b**are	satsuki - **β**are

The input words to *Rendaku* in (48a-c) are pronounced with continuants (*s*, *ʃ*, *h*), as in (48 i). The corresponding output words are often pronounced with non-continuants (*dz*, *dʒ*, *b*) post-nasally, as in (48 ii), and as continuants intervocalically, as in (48 iii). Based on this observation, Section 2.1 concluded that post-nasal fortition plays a role in yielding the output forms in (48 ii). In turn, this ultimately led to the analysis that *Rendaku* applies in *ha-gyoo* to underlying *ɸ* and derives *β*, to which fortition applies and derives *b* post-nasally. For the example in (48c-ii), the relevant derivation would be *nihoɴ-ɸare* → *nihoɴ-βare* → *nihoɴ-**b**are*. Under this analysis, the input for (48c-i), i.e. the underlying form for the word *hare* produced in isolation, would be ***ɸ**are* (rather than ***p**are*).

The results of Experiment 1 suggested that this same fortition process is also responsible for the underlying fricatives *z* and *ʒ* appearing as affricates post-nasally in present-day Japanese loanwords, as in (49-i).

(49)

	(i) Output: **post-nasal underlying voiced <u>non-continuants</u>**			(ii) Output: **intervocalic underlying voiced <u>continuants</u>**		
a.	kaɴ**dz**asu	'Kansas'	(< kænzəs)	juu**z**aa	'user'	(< juzə˞)
b.	paɴ**dʒ**ii	'pansy'	(< pænzi)	bi**ʒ**itaa	'visitor'	(< vɪzətə˞)

The hypothesis that the lenition of voiced non-continuants introduced in medieval Japanese (as exemplified in (43a)) continues to apply in present-day Japanese also helps account for how the voiced non-continuants *dʒ* and *b* in the original Western words were observed to be phonetically realized as continuants (*ʒ*, *β*) intervocalically as in (50-ii) below (in contrast to post-nasal position in (50-i)).[37]

(50)

	(i) Output: **post-nasal underlying voiced <u>non-continuants</u>**			(ii) Output: **intervocalic underlying voiced <u>continuants</u>**		
a.	**dʒ**iɴ**dʒ**aa	'ginger'	(< **dʒ**ɪn**dʒ**ə˞)	maneeʒaa	'manager'	(< mænə**dʒ**ə˞)
b.	naɴ**b**aa	'number'	(< nʌm**b**ə˞)	haa**β**aa	'harbor'	(< haɪ**b**ə˞)

This finding suggests that intervocalic lenition has also played a role in yielding the pronunciation of these loanwords in Japanese. Thus, taken together as a whole, the data in (48)-(50) are evidence that both fortition and lenition of voiced obstruents not only *historically did* but also *currently still do* apply in Japanese, which induce the fortis-lenis alternations of the voiced obstruents examined in these examples. When neither of these

[37] Recall that Japanese loanwords coming from English words with [dz] and [ʒ] are limited and as such were not included in Experiment 1.

positionally conditioned processes apply, *z*/*ʒ* and *dʒ*/*b* in the original loanwords are realized as they are — as continuants in (49-ii) and non-continuants as in (50-i), respectively.[38]

Recall here the comparison between (42) and (43), which was made by the author of *Kenshukuryookoshuu* (蜆縮凉鼓集:1695) in order to promote prescriptive pronunciation in (42). Crucially, the patterns just observed in (48)-(50) above exactly match what was claimed to be the "incorrect" pronunciation in (43). This suggests that the lenition and fortition processes in question, summarized in (51) below, began to take effect in medieval Japanese and have continued to take place up until present-day Japanese.[39]

(51) a. Intervocalic lenition: { dʒ, dz } → { ʒ, z } / V__V
b. Post-nasal fortition: { ʒ, z } → { dʒ, dz } / N__

As illustrated by (44) and (45), these fortition and lenition processes eventually became more disorderly in their application and yielded the "confusion of four *kana* characters." It is proposed here that the 'mix-ups' in pronunciation behind this phenomenon stemmed from some speakers' overextension of the phonological environments for the application of fortition and lenition (beyond only post-nasal and intervocalic positions, respectively). Eventually, this led to a neutralization in terms of the environments where these processes applied.

Moreover, it is also hypothesized that the same neutralization is still at work in present-day Japanese. As a supplement to (51), this hypothesis can help explain: (i) why post-nasal and intervocalic positions contrast in terms of the continuancy of voiced obstruents in present-day Japanese, and (ii) why this positional contrast is not categorically all-or-nothing. Thus, there is independent motivation to postulate that the positional fortis-lenis alternation suggested by the results of Experiment 1 is active in the phonology of present-day Japanese. Crucially, this alternation includes the phonetic realization of the *Rendaku*-voiced *ha-gyoo* sounds as *b*~*β*, as expected under the analysis proposed in the present work.

This "neutralization of environments" has one other aspect that is worth reporting: the second-hand historical data available to us contain virtually no examples of *post-nasal lenition* for the affricates *dʒ*/*dz*.[40] In contrast, examples illustrating the mix-up

---

[38] Note that both post-nasal fortition and intervocalic lenition must be postulated even if one assumes that only one of the fortis-lenis pair in the original loanwords is borrowed into Japanese. For example, even if it is assumed that only /z/ in English (rather than both /dz/ and /z/) is borrowed into Japanese, *kaɴ**dz**asu* (<*Kan**s**as*) in Japanese would have to be captured with fortition. Likewise, even if it is assumed that only /dʒ/ in English (rather than both /dʒ/ and /ʒ/) is borrowed into Japanese, *manee**ʒ**aa* (<*mana**g**er*) need to be captured with lenition.

[39] More specifically, it has generally been assumed that these processes started to apply after the affrication of *d* (in (41)) became established, which made the pronunciations of *dʒi*/*ʒi* and *dzu*/*zu* closer and paved the way to the confusion of the four *kana*s. At the same time, the phonologically-conditioned fortition and lenition processes in (51) most likely were one factor contributing to the confusion of the four *kana*s and, as such, started to apply sometime before complete entanglement arose between the relevant *kana*s and their pronunciation. Alternatively, it is also possible that post-nasal fortition started earlier (before the affrication of *di*/*du* in (41)) and intervocalic lenition started later (after (41) took place).

[40] The only exceptions we are aware of in the philological literature are *naɴ**ʒ**i* (汝) 'you' and *naɴ**ʒ**iɸu* (難渋/難重) 'difficulty' cited by Hashimoto (1966:107) from the 1622 tale *Mikawa-monogatari*. However, in the latter, the prescriptively incorrect Chinese character 重 actually starts with underlying *ʒ* rather than

between the fricatives *ʒ*/*z* and affricates *dʒ*/*dz* are well-attested in other positions. This discrepancy suggests that while the application of fortition and lenition were generally mixed-up in medieval Japanese, the only exception may have been that lenition was still strongly discouraged from applying post-nasally (whereas at the same time, fortition applied in the same environment quite strictly). Section 4.2 below addresses how the neutralization of environments in question has become even more thorough in present-day Japanese.[41]

The results of Experiment 2 (in Section 2.2) provide an interesting perspective on just how broad the scope is for these fortis-lenis alternations. Recall how that experiment examined the ways in which the voiced stops *d* and *g* are phonetically realized post-nasally and intervocalically. The results are illustrated in (52) and (53) below. The input stops in (52) are *Rendaku*-voiced, while those in (53) are inherently voiced in the original Western loanwords.

(52)	(i) Output: **post-nasal *Rendaku*-voiced <u>non-continuants</u>**	(ii) Output: **intervocalic *Rendaku*-voiced <u>continuants</u>**
a.	paɴ-**d**ane 'bread dough' (< **t**ane 'seed')	mono-**ð**ane 'source of anything'
b.	ʃiɴ-**g**ao 'new face' (< **k**ao 'face')	ɸuɾu-**ɣ**ao 'old face'

(53)	(i) Output: **post-nasal underlyingly voiced <u>non-continuants</u>**	(ii) Output: **intervocalic underlyingly voiced <u>continuants</u>**
a.	ɾateɴ-**d**aɴsu 'Latin dance'	ɸooku-**ð**aɴsu 'folk dance' (< **d**æns)
b.	metaɴ-**g**asu 'methane gas'	haiki-**ɣ**asu 'exhaust gas' (< **g**æs)

As shown in (52-i), the inputs to *Rendaku voicing* in (52a-b) are non-continuants (*t* and *k*) but their outputs are predominantly pronounced as continuants (*ð* and *ɣ*) intervocalically as in (52-ii). The results are essentially identical when the inputs are underlyingly voiced — the target obstruents in (53-ii) are predominantly pronounced as continuants (*ð* and *ɣ*) despite the fact that they correspond to non-continuants in the original Western words (*d*, *g*). (The non-continuancy of the input words is retained post-nasally in both (52-i) and (53-i).) Thus, on the whole, the instances of *d*→*ð* and *g*→*ɣ* lenition suggested by these results is directly parallel to the instances of *dʒ*→*ʒ* and *b*→*β* lenition discussed above in (50). This suggests that fortis-lenis alternations in question are quite broad in scope, cross-cutting a broad range of places of articulation.

dʒ, making it unclear whether the word involves post-nasal lenition after all. In fact, *Mikawa-monogatari* is known to contain numeous other spelling mistakes and orthographic discrepancies, extending well beyond the scope of just the confusion of the four *kana*s.

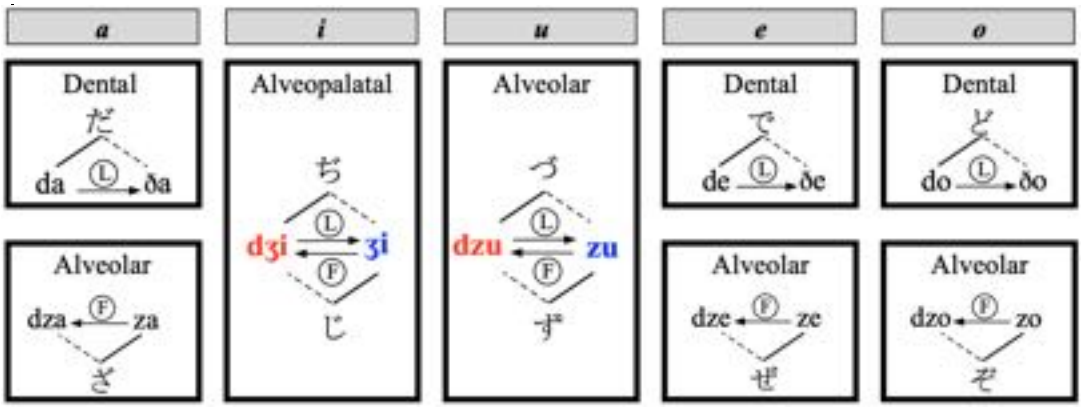
[41] If the alternative chronology of fortis-lenis alternation mentioned at the end of footnote 39 turns out to be correct, this 'rejection' of post-nasal lenition in medieval Japanese receives a natural explanation: that is, relative to intervocalic lenition, post-nasal fortition started spreading earlier, remained in people's grammar longer, and became more stable.

When the present study's experiment on present-day Japanese is connected to the philological literature on diachronic changes in Japanese phonology, as just done in the preceding paragraphs, it not only yields new insights regarding the nature of the "confusion of the four *kana*s" itself but also (and perhaps even more importantly) reveals something that seems to have gone unnoticed *behind* this phenomenon. Namely, the fortition and lenition processes leading to the mix-up of pronunciation have been operative on a much larger scale than has generally been considered, going beyond the "confusion of the four *kana*s" alone (which involves only *dʒi*/*ʒi and dzu*/*zu*) in three different ways. First, these processes apply without distinction to all voiced continuants (*z*, *ʒ*, *β*) and all voiced non-continuants (*dʒ*, *dz*, *b*, *d*, *g*) in Japanese, as was just seen above. Second, their application is not limited to context where the input obstruent precedes *i* or *u* but, rather, can apply in front of any vowel. Note, for instance, how lenition and fortition both apply before *a* in the historical data in (43a-i) and (43b-i). Third, these processes apply regardless of whether the obstruent in question is morpheme-initial or morpheme-internal (cf., for example, (48-ii) vs. (49-i) and (52-ii) vs. (50-ii)).

4.2. Partiality of pronunciation-*kana* entanglement:

The picture just sketched raises two new questions to be answered. First, considering how the entanglement in the pronunciation of *dʒi*/*ʒi and dzu*/*zu* caused the confusion of the four *kana*s (ぢ/じ and づ/ず), why did the same phenomenon not occur for the other three vowels (*a*, *e*, and *o*) and thereby cause similar confusions for the *kana* pairs だ/ざ (*da*/*za*), で/ぜ (*de*/*ze*) and ど/ぞ (*do*/*zo*)? The key to answering this question lies in comparing the expected consequences of the entanglement for ぢ/じ and づ/ず as opposed to だ/ざ, で/ぜ and ど/ぞ. The relevant facts are summarized in (54) below, which contains every *kana* in *da-gyoo* (upper half) and *za-gyoo* (lower half) for all five vowels.

(54)



Inside each column (one for each vowel), there is a box for each place of articulation (dental, alveolar, or alveopalatal) that is possible for a *da*- or *za-gyoo* obstruent in that vowel context. Each such box contains the relevant *kana*(s) and the range of possible pronunciations for each. Horizontal arrows indicate lenition (Ⓛ, pointing rightward) or fortition (Ⓕ, pointing leftward). Solid diagonal lines indicate each *kana*'s original pronunciation-orthography mapping, and dashed diagonal lines indicate an extended mapping (as the result of fortition or lenition).

Crucially, in the columns for *i* and *u*, the output of lenition is identical to the input of fortition (i.e., *ʒi* and *zu*) and, at the same time, the output of fortition is identical to the input of lenition (i.e. *dʒi* and *dzu*). This is precisely the state of affairs that gave rise to the accidentally identical sounds for the four *kanas* in question (ぢ/じ and づ/ず). Crucially, however, this *only* holds true for these four *kana*s. For the remaining pairs (だ/ざ, で/ぜ and ど/ぞ), the obstruents in question always remain distinct. For instance, the output of lenition applying to the dental obstruent in *da* yields a *dental* approximant in *ða*, which is distinct from the *alveolar* fricative in *za*, the 'original' pronunciation of ざ. Similarly, the application of fortition to *za* (ざ) yields an *alveolar* affricate in *dza*, which is distinct from the *dental* stop in *da* (だ). The same can also be said for *de*/*ze* (で/ぜ) and *do*/*zo* (ど/ぞ**)**. This provides the answer to the question why the confusion of *kana*s arose only with two vowels *i and u* while the mix-up of pronunciation involved all five vowels.

The second question to be answered is why something similar to the confusion of the four *kana*s did not arise for the other pairs of voiced obstruents, namely *g*/*ɣ* and *b*/*β*, irrespective of the vowel following them. The answer in fact is straightforward: each of these two pairs involves only a single column of *kana*s (がぎぐげご and ばびぶべぼ, respectively). Thus, even though a similar "mix-up" in continuancy has indeed taken place (as verified by the experiments in the present study), there exists no room for entanglement to arise in the relationship between the relevant *kana*s and their pronunciations. This is presumably why, in previous studies (both diachronic and synchronic), the allophonic variation within the pairs *d*/*ð*, *g*/*ɣ*, and *b*/*β* has not been linked to the fortition and lenition affecting z/*ʒ* and *dz*/*dʒ* in the confusion of the four *kana*s (and, thus, the two have been treated as separate topics).

4.3. Dominance of *za-gyoo kana*:

If correct, the analyses laid out above offers a possible solution to a puzzling asymmetry observed in the results of the two phonetic experiments reported in Section 2. Recall from the last part of Section 2.3 that *z*/*dz* and *ʒ*/*dʒ* exhibit a significantly lower percentage of bursts than *β*/*b*, *ð*/*d*, and *ɣ*/*g* in (30) and (31), as indicated by the differences in percentages above vs. below the thick horizontal lines. This suggests that, even when the "mixed-up" application of fortition and lenition is taken into consideration, *z*/*dz* and *ʒ*/*dʒ* are realized as fricatives much more frequently than would otherwise be expected.

One cause of this asymmetry may be purely phonetic in nature, reflecting the contrast between fricatives (*z*, *ʒ*) and approximants (*β*, *ð*, *ɣ*) in (30) and between affricates (*dz*, *dʒ*) and stops (*b*, *d*, *g*) in (31). This explanation, however, fails to account for why *dz*/*dʒ* involve so few bursts even post-nasally. Rather than a phonetic factor of this sort, it is also possible that this asymmetry may be partially traceable to the inclusion of *Gendai Kanazukai* 'Modern *kana* usage' into school curriculum and the resultant influence on speakers' daily use of the Japanese orthography. Recall from Section 3.3 that this policy recommended that, with certain exceptions, the *kana*s representing the fricatives *ʒi* (じ) and *zu* (ず) be extended to the affricates *dʒi* and *dzu*, thereby largely replacing ぢ and づ, reflecting the fact that *ʒi* and *zu* were already becoming increasingly frequent in some words where *dʒi* and *dzu* were expected. This change in the writing system has encouraged words that originally had *dʒi* and *dzu* as in (55) below to be written with じ

and ず, both intervocalically ((55a-b)) and post-nasally ((55c-d)).[42]

(55)	a.	湯治	とうじ	tou**ʒi**	'hot-spring cure'	(<とうぢ tou**dʒi**)
	b.	稲妻	いなずま	ina**zu**ma	'lightning'	(<いなづま ina**dzu**ma)
	c.	汝	なんじ	nan**ʒi**	'thou'	(<なんぢ nan**dʒi**)
	d	心電図	しんでんず	ʃinden**zu**	'cardiogram'	(<しんでんづ ʃinden**dzu**)

The effect of this orthography reform was not limited to the contexts preceding the vowels *i* and *u*. All voiced obstruents transcribed with *za-gyoo kana*s regardless of the following vowel (i.e., even before *a*, *e*, and *o*) also underwent a similar transition, as exemplified below in (56a-d) for intervocalic position and (56e-f) for post-nasal position.[43]

(56)	a.	*sore**ʒa**a*	それじゃあ	'then (informal)'	(< *sore**dʒa**a* それぢゃあ < *sore**de**-wa* それでは)
	b.	*ʃuu**ʒa**ku*	しゅうじゃく	'obsession'	(< *ʃuu**dʒa**ku* しゅうぢゃく)
	c.	*ee**ʒe**nto*	エージェント	'agent'	(< English *eɪ**dʒ**ənt*)
	d.	*kano**ʒo***	かのじょ	'she/girlfriend'	(< *kano**dʒo*** かのぢょ)
	e.	*en**ʒe**ru*	エンジェル	'angel'	(< English *eɪn**dʒ**əl*)
	f.	*kan**ʒo**o*	かんじょう	'calculation/bill'	(< *kan**dʒo**o* かんぢょう)

When such a transition in orthographic practices is observed in the society, it is not too difficult to imagine that it may also have gradually induced the transition of pronounciation from *dʒ* and *dz* to *ʒ* and *z*.

The results of the present study's experiments suggest that this overwhelming tendency toward lenited pronunciations has also been extended to cases like (57), in which *dʒ* and *dz* are expected as the output of post-nasal fortition.

(57)	a.	かんざくら	kan**ʒa**kura	rather than	kan**dza**kura	'winter cherry blossoms'
	b.	とんじる	ton**ʒi**ru	rather than	ton**dʒi**ru	'miso soup with pork'
	c.	ぽんず	pon**zu**	rather than	pon**dzu**	'soy sauce with citrus juice'
	d.	みんみんぜみ	min**mi**n**ze**mi	rather than	min**mi**n**dze**mi	'robust cicada'
	e.	どんぞこ	don**zo**ko	rather than	don**dzo**ko	'absolute bottom'

[42] In fact, nowadays, even one class of exceptions (words whose etymology-based *kana* transcription of *dʒi* and *dzu* as ぢ and づ were accepted in *Gendai Kanazukai* as an established convention) have begun to be written with じ and ず. It is not difficult to find (especially on the Internet) examples such as みかずき 'crescent' (< ***ts**uki* 'moon') and はなじ 'nosebleed' (< ***tʃ**i* 'blood'), reflecting these words' actual pronunciations with *ʒi* and *zu*. The typing of Japanese characters in the Roman alphabet on computers and mobile devices may also have been one of the factors behind this change.

[43] All words in (55) and (56) (and those in (i) in footnote 44 below) are listed with the indicated *kana* spelling in *Kojien* (5th edition) and *Meikyo-Kokugo-Jiten*, both of which are widely used Japanese dictionaries. Since underlying *dza*/*dze*/*dzo* does not exist in Japanese, this phenomenon is observed only with *dʒa*/*dʒe*/*dʒo*. Moreover, instances with *dʒe* are limited to loanwords (except in some dialects).

In the experimental results in (30) from Section 2.3, surface *z*/*dz* exhibited an extremely low percentage of bursts in intervocalic position (regardless of the following vowel). Moreover, in the first two rows of (31), *dʒ*/*dz* exhibited noticeably few bursts even post-nasally. Recall from the historical data in Section 4.1 that, in medieval Japanese, the lenition of *dz*/*dʒ* was strongly discouraged from applying in post-nasal position, whereas fortition was strictly enforced in that position. The overwhelming tendency towards *z* and *ʒ* in present-day Japanese even in post-nasal position implies that these restrictions have been either removed or significantly weakened for many speakers.[44]

Thus, it is conceivable that the implementation of the orthography reform may have been a factor which further accelerated the *dominance* of the *za-gyoo* in pronunciation (and writing), and the line of reasoning pursued here (based on the consequences of orthography reform and subsequent spread in contexts for lenis pronunciation) provides a rather natural account of the otherwise puzzling asymmetry between (*z*/*dz*, *ʒ*/*dʒ*) and (*β*/*b*, *ð*/*d*, *ɣ*/*g*) observed in the experimental results in (30) and (31).[45]

4.4. Distribution and derivation of *p*:

This section addresses three topics that have remained unaddressed thus far: (i) how the *p* sound is distributed within the Japanese vocabulary, (ii) the derivational status of *p* in relation to the *ha-gyoo* consonants from a phonological point of view, and (iii) how its relevant diachronic changes have come to be reflected in present-day Japanese.

Recall from Section 3.2 that the *ha-gyoo* consonants have followed a rather winding path over the course of their evolution. First, they all originated from the independent phoneme *p* in the ancient period, apparently with no restriction as to its position of occurrence within a word. Then, during (or before) the medieval period, *p* came to be replaced by *ɸ* in most of the vocabulary (everywhere except for some onomatopoeia and loanwords). What complicated the situation further was that *p* started to re-emerge again in limited contexts while it was still being replaced by *ɸ* elsewhere. More specifically, as was briefly described in Section 3.2, during the medieval period, *p* gradually replaced *ɸ* in native Japanese words and Sino-Japanese words as in (58a) and (58b), respectively.

(58) a. Native-Japanese words:

				First attestation
ap**p**aɾe	(あっぱれ)	'admirable'	cf. *aɸ**p**aɾe* (あはれ)	1219
ya**pp**aɾi	(矢っ張り)	'as expected'	cf. *jaɸ**p**aɾi* (やはり)	1477

[44] The overwhelming tendency towards lenition even in post-nasal position can also be observed in the sequence *dʒu* (ぢゅ), which is transcribed as じゅ and frequently pronounced as *ʒu* in present-day Japanese, as illustrated in (i).

(i) a. ねんじゅう neɴ**ʒu**u 'all the time' (< ねんぢゅう neɴ**dʒu**u)
b. しんじゅう ʃiɴ**ʒu**u 'love suicide' (< しんぢゅう ʃiɴ**dʒu**u)

[45] If distinct groups of speakers are influenced by *Gendai Kanazukai* in distinct degrees, that may also cause variability among speakers in the dominance of the *za-gyoo* pronunciation. This possibly underlies the considerable inter-individual variation in the rate of affricate articulation reported in the corpus study by Maekawa (2010a:371).

b. Sino-Japanese words:

kiɴ**p**eɴ	(近辺)	'neighborhood'	cf. ***ɸ***eɴ (辺)	1604~8
ɾip**p**uku	(立腹)	'getting angry'	cf. ***ɸ***uku (腹)	1776

The phonological processes yielding *p* in (58a-b), however, applied only in a limited context — namely, after syllables closed by a moraic consonant. Due to the avid importation of Chinese words, medieval Japanese gradually came to accept such syllables, which fall into the two categories listed in (59) (Hamada 1960:25). (See also Kubozono (1999:217-229) for relevant discussion.)

(59) a. [$_{\sigma}$ C$^{(j)}$V**N**] ___ (N = moraic nasal)

b. [$_{\sigma}$ C$^{(j)}$V**Q**] ___ (Q = moraic obstruent serving as the initial consonant in a geminate)

Closed syllables of this kind ultimately gave rise to the phonological process that reintroduced *p* into the grammar of medieval Japanese.[46] Moreover, the fact that the words in (58a-b) still are actively used today suggests that this same phonological process is still maintained in present-day Japanese. The synchronic derivations for these words in present-day Japanese are presumably as in (60).

(60)

(i)	kiɴ- **ɸ**eɴ				→	kiɴ**p**eɴ		'neighborhood'
(ii)	ɾitu-**ɸ**uku	→	ɾit_-**ɸ**uku	→	ɾi**Q**-**ɸ**uku	→	ɾi**Q**-**p**uku	→ ɾi**pp**uku 'getting angry'
(iii)	ja**ɸ**ari			→	ja**Q**ɸari	→	jaQ**p**ari	→ ja**pp**ari 'as expected'

In (60-i), ɸ→p applies after a syllable closed by a moraic nasal as in (59a). In (60-ii), the closed syllable arises when the final vowel in the preceding open syllable is deleted.[47] In (60-iii), the closed syllable is created through the insertion of an "emphatic" underspecified moraic obstruent (Q).[48] Additional examples of Native and Sino-Japanese words from this period containing *p* are listed in (61) below along with their written attestation dates.

---

[46] Komatsu (1981:275-6, 279) speculates that the appearance of *p* after a closed syllable in Sino-Japanese words performed the same function as *Rendaku* voicing in Native words — namely signaling compoundhood — while still allowing the speakers to deduce the original form of the second member of the compound. He also notes that, since voicing was contrastive for obstruents in Chinese, *Rendaku* would not have been a possible option to perform this function, hence the need for *p* as a new "marker." See also Takayama (1994:351), who argues that a similar kind of cohesion between morphemes is achieved by emphatic gemination.

[47] In the second step in (60-ii), the syllable-final *t* is assumed to lose its phonological features and turn into an underspecified obstruent Q, to which the features of *p* are spread in the final step. This particular derivational analysis, however, is not essential to the point being made here. See Itô (1988) for relevant discussion.

[48] Gemination can also be induced by vowel deletion in native Japanese words, e.g., *kaki-ɸarau*→ *kapparau* 'snatch' and *yowi-ɸarau*→*yopparau* 'get drunk'. (See (69a) below for a detailed derivation of the former.) Regardless of whether the source is such cases of vowel deletion or the insertion of Q, gemination in native Japanese words carries a distinct meaning, signaling emphasis, vividness, and/or informality without altering the lexical meaning of the input word. This is unlike Sino-Japanese words, where the gemination arising from *ɸ*→*p* does not seem to produce any such emphatic effect.

(61) a. **Native Japanese**:

ja**pp**aɾi	(矢っ張り)	'still/as expected':	1477, 1650,1696, 1786, 1809~13, 1899
kata**pp**aʃi	(片っ端)	'one after another':	1713, 1751, 1776~1801, 1808, 1870,~76, 1900
ka**pp**aɾau	(掻っ払う)	'snatch':	1816~26, 1926, 1928
kaɾa**pp**eta	(空っ下手)	'quite unskilled':	1884, 1900, 1954

b. **Sino-Japanese**:

se**pp**uku	(切腹)	'*harakiri* suicide':	1665, 1763~9, 1809~13
ʃiɴ**p**ai	(心配)	'anxiety':	1826, 1832~3, 1857, 1885~6, 1914
saɴ**p**atsu	(散髪)	'hair cut':	1867, 1871, 1921
ze**pp**iɴ	(絶品)	'rarity':	1875, 1929

Recall from Section 3.2 that the change of *p* into *ɸ* most likely started as early as 8C on a small scale in the west and gradually spread toward the east and became more pervasive by 15~16C. In this light, the earliest attestation dates of the examples in (61a-b) — as early as 15C — suggest that the re-emergence of *p* in the contexts discussed above took place either during or after the change of *p* into *ɸ*. Furthermore, all of the examples in (58) through (61) are present in the vocabulary of present-day Japanese and are actively used with the indicated pronunciation. This fact suggests that the relevant phonological changes have been in effect from medieval times (during or after the *p*→*ɸ* change) up to the present.

Unlike the examples in (58) and (61), in onomatopoetic morphemes, *p* freely appears in a broad range of phonotactic environments, e.g., word-medially after an open syllable or word-initially, as illustrated in (62a-c) below.

(62) a. **p**akuɾi / **p**aku**p**aku 'snapping/biting'
b. **p**ikaɾi / **p**ika**p**ika 'flashing'
c. **p**ukaɾi / **p**uka**p**uka 'puffing/floating'

The *p* sound appeared in such words since ancient Japanese, and this pronunciation has been maintained throughout history (Toyama 1972:197, Komatsu 1981:269, Kubozono 1999:62). Likewise, *p* may also appear almost anywhere in loanwords, reflecting the diversity of positions where *p* can appear in the original words from the source languages:

(63) a. ***p****aɴ* 'bread' (< Portuguese *pão*)
b. *su****p****uuɴ* 'spoon' (< English *spoon*)
c. *ʃaɴ****p****aɴ* 'champagne' (< French *champagne*)
d. *kop****p****u* 'glass/cup' (< Portuguese *copo* and Dutch *kop*)

While *p* may appear word-initially in some purportedly native words as in (64) below, these words are arguably derived from onomatopoeia or loanwords.

(64) a. ***p****akuɾi* 'plagiarizing' < *pakuɾi* 'sound of gobbling' (onomatopoeia)
b. ***p****iɴ-hane* 'kickback' < *pinta* 'dot' in Portuguese (loanword)
c. ***p****eɴ****p****eɴ-gusa* 'shepherd's purse' < *peɴpeɴ* 'sound of Japanese banjo' (onomatopoeia)

Quite importantly to the research on *ha-gyoo* consonants in the present work, the diachronic and synchronic examinations above urge us to adopt the following view. Excluding loanwords and onomatopoeia, all *p* sounds in present-day Japanese are derived from instances of *ɸ* occurring in the contexts listed in (59), i.e. after a closed syllable. As such, the appearance of *p* faithfully reflects the segregation established when *p* was reintroduced into the Japanese vocabulary — that is, native and Sino-Japanese words have *ɸ* everywhere except after a closed syllable, whereas loanwords and onomatopoeia can freely have either *ɸ* or *p*.[49]

Unlike the phonologically derived *p* just discussed, *p* as an independent *phoneme* in ancient Japanese apparently had no restriction on the positions where it could occur within a word. After the *p* phoneme was replaced with *ɸ* during the medieval period, the same was also true for *ɸ*, *mutatis mutandis*.[50] This free distribution in *ha-gyoo* sounds (except after a closed syllable) is inherited in present-day native and Sino-Japanese words, as described in Section 3.2.

4.5. Motivation for the derived *p*:

As introduced in Section 1.1, McCawley (1968) proposed that *p* is the underlying phoneme behind all *ha-gyoo* sounds in present-day Japanese. Moreover, Frellesvig (2010:311, 313-4) claims that the historical change during 13~17C (described above as a *p*→*ɸ* change) was as described in (65) below. Note in particular that, under Frellesvig's analysis, instances of *p* following a closed syllable (as in (65a)) are treated as realizations of underlying *p* throughout history (i.e. from ancient to present-day Japanese).

(65) a. /p/ > /p/ {N, Q} __
b. /f/ elsewhere (*f* "possibly realized as [ɸ]" (p. 204))

While this seems to be the standard analysis adopted in the literature, it has at least four problems, each of which will be described below in turn. Taken together, these four problems make it critical to pursue alternative explanations such as the one proposed here (that *p* was derived from *ɸ* after a closed syllable in the medieval period and that such a derivation is still at work in present-day Japanese).

First, after some Western words were loaned into Japanese, instances of *f* in post-nasal position subsequently underwent fortition to *p*, such as in the following examples. (Each word's date of first written attestation is provided in parentheses.)

(66) a. kon**p**eitoo < Portuguese con**f**eito 'candy' (1625)
b. ɾin**p**a < Dutch lym**ph**a 'lymph' (1872)[51]

---

[49] Note that the the historical replacement of *p* with *ɸ* is a case of lenition, whereas the re-emergence of *p* can be thought of as a case of fortition. It then may not be entirely out of the question to speculate that the co-occurrence of these two processes may have served as partial triggers for the fortition and lenition underlying the "confusion of the four *kana*s" — a possibility which may be worth examining in the future.

[50] Recall that, for *ɸ*, one further phonological change *ɸ*→*w* (*ha-gyoo tenko*) took place in word-medial positions (Footnote 21, Section 3.2.1).

[51] Irwin (2001:39) lists *lym**f**e* as its Dutch source. *Nihon Kokugo Daijiten* does not attest its pronunciation at this date (1872), transcribing it with Chinese characters 列印巴, although it is pronounced with *p* in present-day Japanese. It may in fact have been loaned from German *lym**ph**e* (Irwin 2001:40).

The original pronunciation with *ƒ* in (66a) was retained as a fricative in all of its written attestations in 1625, 1688 and 1874~76. Only later did it come to be pronounced as *p* (as first attested in writing in 1912), and this pronunciation has been maintained until the present time. In contrast, a similar loanword, in (67) below, has retained its pronunciation with a fricative throughout all attestation dates up to the present day.

			Written attestations
(67)	aɾu**h**eitou (or *aɾi**h**eitou*)	< Portuguese al**f**eloa 'sugar candy'	1625,1682,1874,1920

Note that the original fricative has come to be fortified to *p in post-nasal position* in (66a-b) while it has been remained to be lenis *in intervocalic position* in (67). The "underlying *p* analysis" in (65) does not offer any insight into this contrast.[52]

Second, as argued by Hamada (1954:24-25), the same fortition that derives *p* from *ɸ* after a moraic obstruent (Q) also seems to be responsible for deriving *ts* from underlying *s*, and *tʃ* from underlying *ʃ* in the cases of 'emphatic gemination' in (68) below. All of these words can still be used in present-day Japanese in informal speech and/or in some eastern dialects.

(68) *s*/*ʃ*→*ts*/*tʃ* in 'emphatic gemination' Written attentations

a.	ot**tʃ**inu	(押っ死ぬ)	'kick the bucket'	cf. **ʃ**inu	'die':	1809-13, 1874-6
b.	det**tʃ**iɾi	(出っ尻)	'big butt'	cf. **ʃ**iɾi	'buttocks':	1809, 1874-6
c.	ma**tts**ugu	(真っ直)	'truly straight'	cf. **s**ugu	'straight':	1870-6, 1964
d.	oto**tts**aɴ	(お父っつぁん)	'dad'	cf. **s**aɴ	'honorable ~':	1809-13, 1836, 1915 1917

Observe, for instance, the parallel derivations between *ɸ→p* in (69a) and *ʃ→tʃ* in (69b).

(69) a. kaki-**ɸ**aɾau → kak_-**ɸ**aɾau → ka-**Q**-**ɸ**aɾau → ka**pp**aɾau (掻っ払う) 'snatch'
b. oti-**ʃ**inu → ot_-**ʃ**inu → o**Q**-**ʃ**inu → o**ttʃ**inu (押っ死ぬ) 'kick the bucket'

In other words, *ɸ→p* in native Japanese words can be regarded as merely one instance of a more general process of adding occlusion to signal emphasis/vividness. With the "underlying *p* analysis," the parallelism between these instantiations of emphatic gemination would have to be regarded as accidental.

Third, in a fair number of words where *p* re-emerged as geminates in medieval Japanese, *p* and *ɸ* are both permitted. Examples of such Sino-Japanese words and native-Japanese words are listed below in (70) and (71), respectively. (Once again, the attestation dates here are limited to citations in *Nihon Kokugo Daijiten*, relying on the pronunciations indicated there via *kana* phonetic guides for Chinese characters (e.g., 立腹[リッフク]) and/or spellings using the Roman alphabet (e.g., Tappitna [タッピツナ]). Allophonic variations within the *ha-gyoo* sounds (e.g. *h* vs. ç) are not indicated here.)

[52] While it could be stipulated that the underlying *p* analysis is restricted to only native and Sino-Japanese words, that would still leave it unaccounted for why *ƒ* changed to *p* in these loanwords. Of course, an *ad hoc* process of post-nasal fortition operating only in loanwords could be posited to account for (66a), but that would incorrectly derive numerous unattested forms like **inpoomaru* (< English 'informal').

					Written attestations
(70)	a.	'getting angry'	ɾi**ɸ-ɸ**uku	(りっ<u>ふ</u>く)	1305, ca. 17C
			ɾi**p-p**uku	(りっ<u>ぷ</u>く)	1776
	b.	'good handwriting'	ta**ɸ-ɸ**itsu	(たっ<u>ひ</u>つ)	1474
			ta**p-p**itsu	(たっ<u>ぴ</u>つ)	1603~4, 1885~6
	c.	'superb piece'	ze**ɸ-ɸ**iɴ	(ぜっ<u>ひ</u>ん)	1474
			ze**p-p**iɴ	(ぜっ<u>ぴ</u>ん)	1875, 1929
(71)	a.	'stretch/pull'	**ɸ**i**ɸɸ**aɾu	(ひっ<u>は</u>る)	early 13C, 1458~60, 1642, ca.1683
			ɸi**pp**aɾu	(ひっ<u>ぱ</u>る)	late 14C, 1698, 1706, 1718, 1803, 1809~13, 1813, 1898, 1911, … 1925
	b.	'entirely (refused)'	ma**ɸɸ**iɾa	(まっ<u>ひ</u>ら)	1471~73
			ma**pp**iɾa	(まっ<u>ぴ</u>ら)	1595, 1639~40, 1748, 1806, 1820~49, 1914, 1952
	c.	'plentifully'	ta**ɸɸ**uɾi	(たっ<u>ふ</u>り)	1586, 1668
			ta**pp**uɾi	(たっ<u>ぷ</u>り)	1708, 1720, 1806, 1809~13, 1811, 1857~63, 1870~76, …, 1909, 1950, 1975

While the variants with *ɸ* in all of these examples are attested earlier than those with *p*, the variants with *ɸ* all disappeared by around 18C. In contrast, the variants with *p* remain in present-day Japanese. The overlap between the range of attestation dates for the *p* and *ɸ* variants in each pair implies that *p* and *ɸ* were in free variation after a closed syllable for several centuries (during which the application of the phonological change *ɸ→p* was optional).[53] Then, when the *ɸ→p* change started to apply more strictly and pervasively around 17~18C, *p* became a phonologically conditioned allophonic variant of *ɸ* in the position after a closed syllable. Under the traditional analysis, whereby underlying /*p*/ is assumed to have simply remained as-is in this position, it is difficult to explain (i) why free variation arose between *p* and *ɸ* in the first place, (ii) why the earliest attestation dates for the derived sound *ɸ* are earlier than those for the underlying sound *p* in so many pairs (precisely the opposite of what the traditional analysis would predict), and (iii) why, later in the history of Japanese, the underlying sound *p* remained while the derived sound *ɸ* tended to disappear.[54]

---

[53] Hamada (1954:25-27) also points out this possibility based upon the coexistence of *nippoɴ, niɸoɴ,* and *niɸɸoɴ* in various Western and Korean records published in 17C. Asada (2000:833), on the other hand, reports that evidence of allophonic variation between *p* and *ɸ* in a post-nasal position can be found as early as in *Shoomyoo Shiryoo* (Japanese Buddhist Hymns) in 13C.

[54] The disappearance of the forms with *ɸ* cannot be attributed to the rule in (65b) simply 'dying out' since that would mean all *ha-gyoo* sounds would need to be pronounced as *p* everywhere. Moreover, since the traditional analysis claims that the *p→ɸ* change applied anywhere *other than post-N/Q position*, even if the rule in (65b) had applied only optionally, it would not have affected sounds in post-N/Q position, including the words in (70) and (71). If, on the other hand, *p→ɸ* is claimed to apply everywhere (including post-N/Q position), its optionality should have allowed free variation between *p and ɸ* everywhere, which is not the case. Thus, variable rule application also cannot be a valid explanation for the post-Q variation observed in (70) and (71).

Finally, post-nasal fortition is a ubiquitous phonological process that has been observed in many languages. For instance, in German and English, a so-called intrusive stop may appear in post-nasal context, as illustrated in (72a) and (72b) (where the fortition in question involves *f*→*pf* and *z*→*dz*, respectively).[55]

(72)	a.	German	*Senf*	'mustard'	[zem**f**]	→	[zem**p**f]
			Hanf	'hemp'	[ham**f**]	→	[ham**p**f]
	b.	English:	*bans*		[bæn**z**]	→	[bæn**d**z]

From this broader cross-linguistic perspective, the process of fortition after a closed syllable proposed above may be regarded as a natural and expected phenomenon. In contrast, the alleged lenition in (65b) does not apply in a similarly natural phonological environment. In particular, note that the environment "elsewhere" in (65b) collapses the heterogeneous environments "V_V" and "#_," rather unlikely bedfellows to constitute a positional condition for phonological changes.

To sum up, the discussion above covered four kinds of data that are problematic for the "underlying *p* analysis" as in (65) but receive a natural explanation under the proposed "underlying *ɸ* analysis": (i) the application of *ɸ*→*p* to older loanwords (e.g., *con**f**eito*→*koɴ**p**eitoo*), (ii) instances of fortition under emphatic gemination with *s*→*ts* and *ʃ*→*tʃ* (e.g., ***ʃ**inu*→*ott**ʃ**inu*) that parallel *ɸ*→*p,* (iii) free variation between *ɸ* and *p* (e.g., *ze**ɸ**-**ɸ**iɴ* ~ *ze**p**-**p**iɴ*), with the former appearing earlier but only the latter remaining to the present day, and (iv) the cross-linguistic ubiquity of post-nasal fortition (cf. English *bans* as [bæn**d**z]). Recall that one of the appeals of McCawley's "underlying *p* analysis" was that it mimics the historical fact that all *ha-gyoo* sounds were pronounced as *p* in ancient Japanese but later came to be replaced by fricatives in most of the vocabulary. However, the four pieces of evidence surveyed above support the idea that the ancient *p* phoneme was replaced by *ɸ* and that, during the medieval period (or earlier), *p* started to re-appear in the main vocabulary due to phonologically conditioned derivation. This makes it difficult to maintain the hypothesis that the grammar of present-day Japanese inherits ancient *p* as the phoneme underlying all *ha-gyoo* obstruents. In other words, despite its tacit appeal for having synchrony mimic diachrony, the traditional analysis is ultimately out of sync with the historical facts.

4.6. Sound-orthography mappings:

In the discussion up to this point, a close investigation of the historical developments in the sound system of Japanese has lent support to the proposed "underlying *ɸ* analysis" of the *ha-gyoo* consonants in present-day Japanese. This section presents additional support for the underlying *ɸ* analysis based on an examination of the mappings between sounds and orthography. In the study of Western languages, we have often witnessed the difficulty of detecting clear and systematic relationships between sounds and their Roman transcription. As a consequence, some researchers are understandably skeptical about the idea that there can be regular and phonologically motivated relationships between the graphs of an orthography and the sounds in some languages. When we carefully examine the history of the development of orthography in Japanese, however, we can be reasonably confident that it represents the sound that embodies historical changes in

[55] We would like to thank Isabelle Darcy and Ken de Jong for pointing out these examples to us.

pronunciation.

As briefly mentioned in Section 1.1, the Japanese syllabary includes some diacriticized *kana*s. In particular, if two tilted short lines (゛), called *dakuten* 'turbid dot', are added to a *kana*s representing a mora beginning with a voiceless obstruent, the new *kana* represents the voiced counterpart (e.g. か *ka* + ゛ (*dakuten*) = が *ga*). The *dakuten* diacritic is also applied to *ha-gyoo* ((73a) below), where the diacriticized *kana*s are *ba-gyoo* ((73b)), i.e. *kana*s representing moras beginning with *b*. In addition to *dakuten*, another diacritic exists that is unique to the *ha-gyoo kana*s. Known as *han-dakuten* 'half-turbid dot' (゜), this small circle diacritic creates *pa-gyoo* ((73c)), i.e. *kana*s representing moras beginning with *p*.

(73) a. *ha-gyoo*: は(ha) ひ(çi) ふ(ɸu) へ(he) ほ(ho)
b. *ba-gyoo*: ば(ba) び(bi) ぶ(bu) べ(be) ぼ(bo)
c. *pa-gyoo*: ぱ(pa) ぴ(pi) ぷ(pu) ぺ(pe) ぽ(po)

The subsections below first review the philological literature regarding when and how each of these two diacritic marks arose in Japanese. This is then supplemented with a critical discussion of the bigger picture suggested by the available evidence.

4.6.1. Voicing and *dakuten*:

In present-day Japanese, voicing in obstruents plays two different functions. On the one hand, it may represent a phonemic distinction, as illustrated in (74a) below. On the other hand, it may also represent an allomorphic distinction established by *Rendaku*, as in (74b). In the latter case, since the two allomorphs in question often share the same meaning and the same Chinese characters (e.g., 仮名 for both *kana* and *gana*), the derivational (i.e. input-output) relation between the two allomorphs is readily recognized by Japanese native speakers.

(74) a. Phonemic distinction:

***k**ama* かま 'sickle' vs. ***g**ama* がま 'toad'

b. Allomorphic distinction:

***k**ana* かな *'kana* script' → *hira-**g**ana* ひら**が**な 'cursive *kana* script'

In both cases, the *dakuten* diacritic (゛) indicates the presence of voicing in an obstruent.

The present-day use of *dakuten* is often traced back to a system of diacritics (known as *dakushooten* 'dots for voiced sounds') that was used by Japanese scholars during the mid to late *Heian* period (mid 11C~ early 12C) to indicate both voicing and tone in Chinese words (Komatsu 1981:61-67, Numoto 1993:15, 27). Originally, these diacritics were placed on the left side of the characters (initially *Manyoogana*, later *katakana*) that served as the phonetic guide (*rubi*) to Chinese characters. After a winding history, they came to be indicated in the top right corner of *katakana*s by the *Muromachi* period (14~16C). Finally, by the *Edo* period (17~19C), the *dakuten* (゛) diacritic came to be commonly used on individual *hiragana*s (independent of Chinese characters) to indicate voicing, i.e., as *dakuten* in today's form, location and use (Tsukishima 1964:85, Numoto 1992:83-4).

Recall from Section 3.1 that voicing did not have rigid phonemically-contrastive

function in ancient to early medieval Japanese, and that it was only later (during late 16C to 17C) it became more clearly contrastive (Kiyose 1985:84-85, Hamada 1960:24). Numoto (1993:28) points out that the development of *dakuten* just described reflects the rise of the contrastive nature of voicing in obstruents, implying that this phonological change gave rise to the practical need to indicate voicing consistently on *kana*s.

4.6.2. Occlusion and *han-dakuten*:

The origin of the *han-dakuten* diacritic (゜) used on *pa-gyoo kana*s, on the other hand, is more difficult to pinpoint. The traditional assumption was that *han-dakuten* first appeared in Christian documents written by Westerners in late 16C and later spread to Japanese documents. (See, for example, Okumura (1972:90).) However, Numoto (1990:1-3, 1992:84-87) argues against this view based on two pieces of evidence. First, the diacritic in question (゜) is found in documents written well before the late 16C. Second, in Japanese documents written during the mid *Edo* period (17C~ early 18C), i.e., well after the Christian documents in question, the use of the *p* sound in a *kana* was rarely indicated with a diacritic. Even in cases where it was, diacritics other than *han-dakuten* (゜) were often used.

Instead, Numoto (1990:8-9) argues that the progenitor to the modern *han-dakuten* was a diacritic (also written ˚) that was originally used to transcribe the pronunciation of *p* and other sounds in Chinese. For example, in *Tooin-shiryoo* (literally "*Tang* sound materials," containing *kana* transcriptions of the *Tang* pronunciation of Chinese characters, 17~19C), this diacritic was used on *kana*s to signal the presence of some Chinese sound that could not be captured by the set of existing *kana*s. Later, during 18C, *Zen* priests and other intellectuals extended the use of this diacritic to native Japanese words. Finally, during 19C, its function ultimately became limited to only indicate *p*.

Hamada (1960:26) argues that the establishment of the *pa-gyoo kana*s was the result of the popularization of literacy education, which promoted the knowledge of Chinese characters and use of *kana*s among the general public. Moreover, Numoto (1990:8) claims that three historical events that took place in Japan during 17~19C served as additional factors. First, during this period, there was renewed interest in the Chinese language among Japanese intellectuals, who perceived the *p* sound in Japanese to be identical to the *p* sound in Chinese. Second, *han-dakuten* was used to indicate *p* in *Tooin-shiryoo*, which was widely used for the study of Chinese at the time. Third, during this period, reference works (including *Tooin-shiryoo*) came to be printed and distributed on a massive scale, and their influence gradually spread to literary works as well. This eventually promoted the widespread use of *han-dakuten* by the general public.

In addition to these various language-external factors, Numoto (1992:87) also considers that one additional language-internal factor helped lead to the establishment of the *pa-gyoo kana*s. He claims that, after *ɸ* developed *h* as an allophonic variant, *p* ceased being an allophone of *ɸ* and instead was established as an independent phoneme. This assertion may have some relevance, but it fails to explain why *pa-gyoo kana*s were created from *ha-gyoo kana*s with the addition of *han-dakuten*. Moreover (and more importantly), it also lacks empirical support since, in present-day native and Sino-

Japanese words, *p* and *ɸ* are still in an allophonic relation after a closed syllable.[56]

It is proposed here that the sharp increase in the use of *pa-gyoo kana*s in 18~19C can be attributed to a different language-internal factor — the re-emergence of the *p* sound (and the corresponding increase in usage frequency) within the Japanese vocabulary. In this connection, however, recall the emphatic gemination data in (39) (Section 3.2.2) from *Heeke Monogatari* (1219〜1222), suggesting that *p* started to re-emerge in the Japanese vocabulary as early as in 13C. If so, then why did it take so long for the *pa-gyoo kana*s to become fully established and widely used by the public? Recall from examples (70)-(71) in Section 4.5 that the *ɸ*→*p* change most likely applied optionally and selectively for several centuries until around 17~18C, even after emphatic gemination started to occur. But apparently, this phonological change started to apply more strictly and pervasively after 18C and, as a consequence, *p* is likely to have become regarded as a phonologically conditioned allophonic variant of *ɸ* after a closed syllable. Once this happened, there would have been a need to ensure that *p* is represented in *kana* transcription in a way different from *ɸ* (and its variants *ç* and *h*). The *han-dakuten* convention therefore was a rather natural candidate to fulfill this function by indicating the addition of occlusion only in the former.

There is historical evidence that the function of *han-dakuten* is indeed to indicate occlusion in *ha-gyoo* consonants. Before *han-dakuten* (゜) was used for *p* alone, there was a period during which it was used to transcribe not only *p* but also instances of *ts* appearing after a moraic consonant (*Q* or *ɴ*) in some dialects spoken in Edo. This is illustrated in (75) below with examples from *Ukiyoburo* (1809-13) and *Ukiyodoko* (1813), adapted from Toyama (1972:243) and Numoto (1990:8-9).

(75)	a.	お父さん	おとっさ゜ん	otoQsaɴ	→	otossaɴ	→	otottsaɴ
	b.	小せへ	ちっせ゜へ	tʃiQsee	→	tʃissee	→	tʃittsee
	c.	お吉さん	おきっさ゜ん	okiQsaɴ	→	okissaɴ	→	okittsaɴ
	d.	番(頭)さん	ばんさ゜ん			bansaɴ	→	bantsaɴ
	e.	お天道様	おてんさ゜ま			otensama	→	otentsama

Note that, in these examples, *han-dakuten* (゜) is added to *kana*s with the fricative *s* (either さ *sa* or せ *se*) to signal that the result is pronounced with an affricate (*ts*). This implies that the function of the *han-dakuten* (゜) diacritic at the time was to indicate the addition of occlusion to a voiceless fricative. Crucially, the use of the *han-dakuten* (゜) on the *pa-gyoo kana*s (ぱぴぷぺぽ) could be described in the exact same terms, i.e., it adds occlusion to the voiceless fricative transcribed by the *ha-gyoo kana*s (はひふへほ).

This hypothesis is directly supported by cases of emphatic gemination of *ha-gyoo* sounds into *pa-gyoo* sounds after a closed syllable, as in (76) and (77) below. (See also the examples in (58) and (61) in Section 4.4.) Here, the emphasis is signaled via the addition of occlusion to the geminated fricative, which is represented in orthography by

---

[56] This relation does not necessarily hold in loanwords and in a small handful of Sino-Japanese words, e.g., *sutaɸɸu* 'staff' or *kençi*/*kenpi* (謙卑) 'modesty'. In addition, in Sino-Japanese compounds of the form [Compound [Compound X+Y] +Z], *ha-gyoo* consonants are allowed after a closed syllable in the final morpheme ('Z'), e.g. [[*gimoɴ*] *ɸu*] 'question mark (疑問符), where the morpheme in question occurs after the boundary of the 'base' compound.

adding the *han-dakuten* (゜) diacritic to the *ha-gyoo kana*s.

(76) a. ja**h**aɾi (やはり) 'still/as expect→ jap**p**aɾi (やっぱり) 'still/as expected'
b. a**h**aɾe (あはれ) 'admirable' → ap**p**aɾe (あっぱれ) 'admirable'

(77) a. çiki (ひき) 'looking up' → kubip**p**iki (くびっぴき) 'constant reference'
b. ɸuɾo (ふろ) 'bath' → çitop**p**uɾo (ひとっぷろ) 'a quick bath'

That *han-dakuten* indicates occlusion can be demonstrated in a different way. According to *Nihon Kokugo Daijiten*, the native Japanese word *jo**ɸ**odo* 'appropriate/very much' was derived in a rather unusual way during the medieval period, as described in (78).

(78) a. *joki **ɸ**odo / joi **ɸ**odo* (よきほど/よいほど) → *jo**pp**odo* (よっぽど)
b. *jo**pp**odo* (よっぽど) → *jo**ɸ**odo* (よほど)

First, the phrasal expression *jo(k)i **ɸ**odo* 'appropriate degree' evolved into the single word *jo**pp**odo*, as in (78a). Then, as illustrated in (78b), *jo**ɸ**odo* was derived as the *non-emphatic* counterpart to *jo**pp**odo* — a "*pp*→*ɸ*" change enacted by *reverse* analogy to the "*ɸ*→*pp*" change observed in other non-emphatic vs. emphatic word pairs like *ja**ɸ**aɾi - ja**pp**a*ɾi 'still' as in (76). That is, the *ɸ* in the non-emphatic form (*jo-**ɸ**odo*) derives from *p* in the emphatic form (*jo**pp**odo*). Orthographically, the end result is that the *han-dakuten* diacritic (゜) is *taken away* (ぽ→ほ), indicating the *removal of occlusion* in pronunciation. Thus, this example demonstrates that the *han-dakuten* diacritic (゜) represents occlusion in exactly the opposite way from (75)-(77).

To sum up, the re-emergence of the *p* sound, coupled with the four language-external factors mentioned by Hamada and Numoto, all contributed to the establishment and pervasive use of *pa-gyoo kana*s in 18~19C. Since then, up through to the present time, the combination of the *ha-gyoo kana*s and a *han-dakuten* diacritic has come to consistently represent the *p* sound, i.e., indicating the addition of occlusion to *ha-gyoo* consonants.

4.7. Sounds and orthography of *ha-gyoo*, *ba-gyoo* and *pa-gyoo*:

4.7.1. Comparison of the two analyses:

Recall from (73) in Section 4.6 that all *ha-, ba- and pa-gyoo* consonants are transcribed with the *kana*s for *ha-gyoo* (e.g., は). In addition, the voicing difference between *ha-gyoo* and *ba-gyoo* is indicated by the presence vs. absence of *dakuten* (゛) and the difference between *ha-gyoo* and *pa-gyoo* consonants is indicated by the presence vs. absence of *han-dakuten* (゜). This section compares McCawley's "underlying *p* analysis" to the proposed "underlying *ɸ* analysis," summarized in (79) and (80), respectively below, in terms of how well they capture these relationships between sounds and orthography.

(79) Underlying *p* analysis (McCawley 1968):

/p/ → ɸ → h, ç ↓ b	a. *Rendaku*:	p → b
	b. Lenition:	p → ɸ
	c. Place of articulation changes:	ɸ → { h, ç }

(80) Proposed underlying *ɸ* analysis:

p ← /ɸ/ → h, ç ↓ b ← β	a. *Rendaku*:	ɸ → β
	b. Post-N fortition:	β → b
	c. Post-N/Q fortition:	ɸ → p
	d. Debuccalization/palatalization:	ɸ → { h, ç }

Under McCawley's analysis, the *ha-gyoo* sounds and their *kana*s are derived when lenition applies to underlying *p* as in (81a) below, and *ba-gyoo* sounds are derived when *Rendaku* voicing applies to underlying *p* as in (81b).

(81) a. Lenition: /pa/ → ɸa → [ha] (*change in place of articulation*)
ぱ は

b. *Rendaku* voicing: /pa/ → [ba]
ぱ ば

In (81a), the lenition is encoded in the orthography by removing the *han-dakuten* (゜) from the *kana* with underlying *p*. In (81b), the voicing is encoded by substituting a *han-dakuten* (゜) with a dakuten (゛). Yet if the *dakuten* (゛) were used consistently to indicate only the addition of voicing, then when *Rendaku* voicing applies to underlying *p*, the *kana*s for *p* should be accompanied by *dakuten*, e.g. ぱ゛ — that is, ぱ plus ゛. In reality, however, *ba*, for instance, is transcribed as ば, i.e. without the *han-dakuten* (゜). In fact, none of the predicted transcriptions for *ba-gyoo* sounds as in (82) are permitted.

(82) *{ ぱ゛ ぴ゛ ぷ゛ ぺ゛ ぽ゛ }

Thus, McCawley's analysis has to rely on surface phonetic values alone and simply associate *ha-gyoo kana*s with *ɸ*, *ba-gyoo kana*s with *b*, and *pa-gyoo kana*s with *p*. Such an approach fails to characterize what the *ha-gyoo kana*s represent in any consistent way. (Note, for example, that no explanation is provided for why it is the *ha-gyoo kana*s that are used as the 'base' glyphs for the *pa-* and *ba-gyoo kana*s.) This approach also fails to attach a consistent meaning to the two diacritics. Thus, the analysis in (81) fails to provide a rational explanation for the correspondences between the relevant consonants in (73a-c) and their corresponding *kana* transcriptions.[57]

The proposed "underlying *ɸ* analysis" in (80), on the other hand, straightforwardly captures the use of diacritics on the *ba-gyoo* (e.g., ば) and *pa-gyoo kana*s (e.g., ぱ), as illustrated in (83).

[57] As pointed out to us by Zendo Uwano, the ill-formedness of the transcriptions in (82) could be easily captured by postulating a constraint prohibiting the doubling of diacritics. This approach, however, would still inherit the two major theoretical shortcomings of McCawley's analysis: (i) failing to account for why the *ha-gyoo kana*s (はひふへほ) are used as the 'base' glyphs and (ii) failing to assign a consistent meaning to the two diacritics added to them.

(83) a. *Rendaku* voicing: /ɸa/ → *βa*
は ば

b. Post-N/Q fortition: /ɸa/ → *pa*
は ぱ

Under this analysis, the addition of a diacritic to a *ha-gyoo kana* captures the intuition that some property has been "added" to the sounds it represents. More specifically, the usage of diacritics can be said to follow the simple principles in (84) below.

(84) (i) Use a "plain" *kana* when a voiceless obstruent as the morpheme-initial phoneme surfaces as-is, i.e. without addition of voicing or occlusion:

e.g., しょ**は**ん *ʃoɦaɴ* 'first edition' (初版)

(ii) Add *daku-ten* (゛) to the *kana* in (i) when **voicing** is added to **the original phoneme**:

e.g., ガリ**ば**ん *gariβaɴ* 'mimeograph' (ガリ版)

(iii) Add *han-dakuten* (゜) to the *kana* in (i) when **occlusion** is added to **the original phoneme**:

e.g., しゅっ**ぱ**ん *ʃuppaɴ* 'publication'(出版)

The principle in (84i) designates voiceless obstruents as defaults and transcribes them with plain *kana*s. It may be considered as a historical heritage from the long period during which voicing was not contrastive in Japanese. The principles in (84ii-iii) then prescribe the usage of diacritics that reflect the phonological changes of voiceless consonants, that is, *dakuten* (゛) for the addition of voicing to all voiceless obstruents *k*/*s*/*t*/*h* ((84ii)), and *han-dakuten* (゜) for the addition of occlusion to *h* ((84iii)).

Three points are worth noting about this approach. First, after /ɸ/ undergoes *Rendaku* and yields *β*, as described in (80a), it may further undergo fortition and be realized as *b* post-nasally, as in (80b). However, since the occlusion in this case applied to *β* rather than the "original phoneme" *ɸ*, the principle in (84iii) does not apply and *han-dakuten* (゜) is not added even though occlusion is involved. Second, although (84i-ii) apply to all obstruents, (84iii) is limited to cases involving /ɸ/ as the underlying phoneme, and as such *han-dakuten* (゜) only appears with *ha-gyoo kana*s, at least in present-day Japanese. (Recall from (75) the obsolete usage of *han-dakuten* on *sa-gyoo*, e.g., さ゜, in *Edo* dialects.[58]) Third, the occlusion indicated by *han-dakuten* (゜) can also be phonemically contrastive, though this use is largely limited to loanwords (e.g. ***h****aato* 'heart' vs. ***p****aato* 'part'), onomatopoeia (e.g. ***h****ara****h****ara* 'fluttering' vs. ***p****ara****p****ara* 'pattering'), and clippings (e.g. ***h****afiri* 'running' vs. ***p****afiri* < *tsukai****p*-*p****a****firi* 'errand runner'). The principle in (84iii) can be easily modified to account for such cases with a supplemental clause, e.g. "Add *han-dakuten* (゜) to the *kana* in (i) when occlusion is added to the original phoneme *or when*

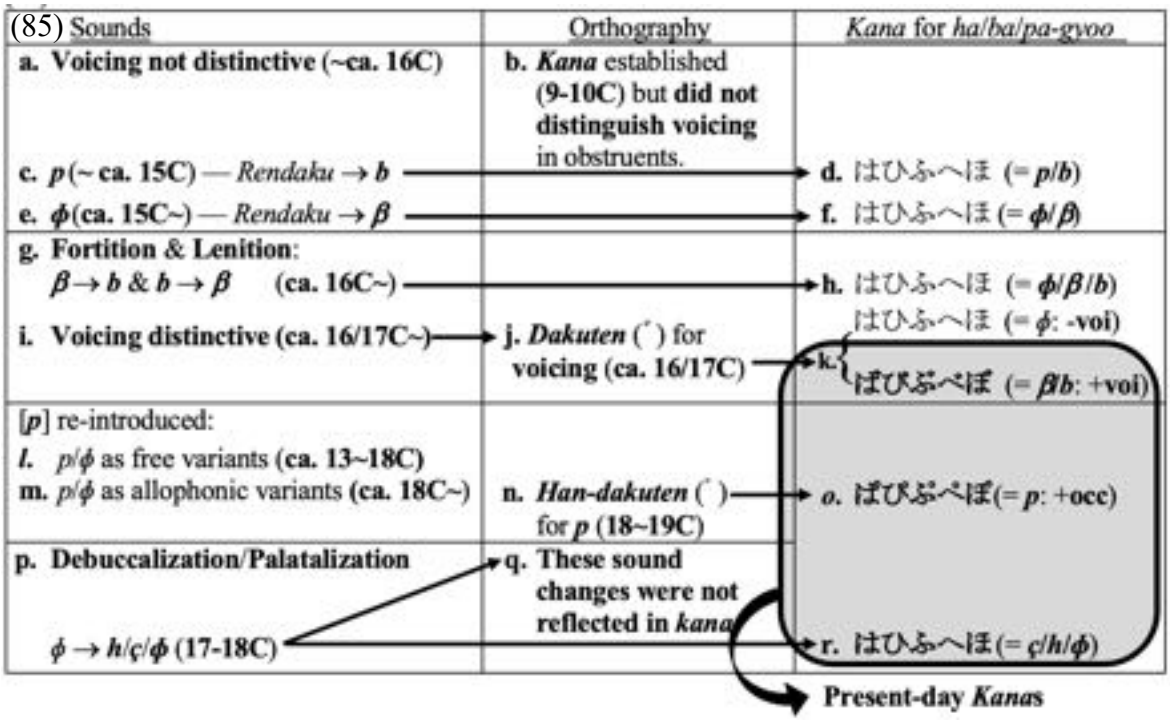
[58] According to Numoto (1990:9), *han-dakuten* eventually ceased to be used on *ts* and *tʃ* for three reasons: (i) *ts* and *tʃ* were less frequent than *p*, (ii) unlike *p*, *ts* and *tʃ* could be represented through other combinations of *kana* (e.g.「ツァ・ツェ・ツォ」), and (iii) by restricting its use to *p* alone, the function of the *han-dakuten* diacritic was simplified.

the occlusivized counterpart of the original phoneme appears as an independent phoneme".

With these three points taken into account, the proposed "underlying ϕ analysis" in (80) depicts the use of *kana* and diacritics for the *ha-gyoo*, *ba-gyoo*, and *pa-gyoo* sounds as a fully rational system. Moreover, describing the diacritics as indicating that some property has been "added" to the relevant sounds may reflect the intuition many native speakers of Japanese have about the *kana* system. In contrast, McCawley's "underlying *p* analysis" in (79) fails to paint a similarly rational picture of sound-orthography correspondences in Japanese.

4.7.2. Overview: Diachrony of sound-orthography correlations

A summary of how the relationship between sound and orthography for *ha-gyoo*, *ba-gyoo* and *pa-gyoo* has changed over time is summarized in (85) below. The left-most column summarizes the relevant historical transitions in sounds, the middle column summarizes those of orthography, and the right-most column summarizes those of *ha-*, *ba-* and *pa-gyoo kana*s. The discussion thereafter follows the order (85 a~r), but note that this does not necessarily proceed in strictly chronological order.



Voicing was not rigidly contrastive in Japanese obstruents, in everyday language, at least until 16C ((85a)). As a result, the inventory of *kana* (*katakana* and *hiragana*), which were established around 9~10C, lacked separate glyphs to mark the voicing distinction ((85b)) (Tsukishima 1964:73, 80, Kiyose 1985:85, Numoto 1992:78, among others).[59]

Originally, all *ha-gyoo* consonants were *p*, and since *Rendaku* voicing applied throughout the history of the Japanese language, it derived *p*→*b* ((85c)). Consequently, the *kana*s はひふへほ represented both *p* and *b* at that time ((85d)). Around 15C, once *p*

[59] Before *kana* was invented, Chinese characters were used to represent the sounds of Japanese in a system called *Manyoo-gana*, which can be traced back at least as far back as 5C.

came to be replaced by *ɸ*, *Rendaku* derived *ɸ*→*β* ((85e)) and the *kana*s はひふへほ represented both *ɸ* and *β* ((85f)).[60]

While *Rendaku* voicing continued to derive *ɸ*→*β* during this period ((85e)), fortition and lenition caused *β* to be realized as *b* post-nasally and *b* to be realized as *β* intervocalically since around 16C ((85g)). These phonological processes as well as the later loosening of their positional conditions started to induce the mix-up of the pronunciation of *β* and *b* in some word-internal positions. During this period, the *kana*s はひふへほ must have represented *ɸ*, *β* and *b* ((85h)).

Voicing became contrastive in obstruents by around 16~17C ((85i)) and came to be consistently indicated with *dakuten* (゛) since around 16C~17C ((85j)) (Tsukishima 1964:85, Numoto 1992:83-4). As a result, the *kana*s ばびぶべぼ came to represent both *β* and *b* (with voicing indicated via *dakuten*), and the *kana*s はひふへほ came to represent only *ɸ* ((85k)).

Since ca. 13C, *p* was gradually reintroduced into the non-onomatopoetic vocabulary. At first, *p* and *ɸ* may have been in free variation ((85*l*)), but by around 18~19C, *p* came to be regarded as an allophone of *ɸ* after a closed syllable ((85m)). The occlusion in *p* came to be consistently marked with *han-dakuten* (゜) ((85n)) (Numoto 1990:8-9), resulting in the representation of *pa-gyoo* with the *kana*s ぱぴぷぺぽ ((85*o*)).

Finally, the allophonic variants of *ɸ* arose around 17-18C — presumably, first *h* due to debuccalization and then *ç* due to palatalization ((85p)). These sound changes, however were not reflected in the *kana*s for *ha-gyoo* (はひふへほ) ((85q-r)), whose five glyphs were already fixed by 10C ((85b)).

Besides those just discussed, no other major changes have taken place affecting the *kana*s or sounds in question (*ha-gyoo*, *ba-gyoo* and *pa-gyoo*).[61] The resulting sound-orthography correspondence in present-day Japanese is indicated in the shaded box, i.e., bottom half of (85k), (85o), and (85r).

---

[60] Recall from Section 3.2, however, that the exact timing of the *p*→ *ɸ* change is not agreed upon within the philological literature.

[61] The only major class of exceptions are the modern-era innovations in (mostly Western) loanwords to approximate phonetic details of the pronunciation of the corresponding word in the source language by creatively combining *katakana*s, e.g., *ɸaɴ* 'fan' (ファン) or *baioɾiɴ* 'violin' (ヴァイオリン). Observing cases like *ɸaɴ* 'fan' (ファン), *ɸiɾumu* 'film' (フィルム), *ɸɛɾuɾo* 'felt' (フェルト), *ɸɔɴdjuu* 'fondue' (フォンデュー), one may question the derivation of *ha-gyoo* consonants from the underlying *ɸ* adopted above (= the end result of (36) inherited to present-day Japanese) and assume that a bilabial fricative and a pharyngeal fricative are different phonemes in contemporary Japanese (because *ɸaɴ* 'fan', for example, is not realized as *haɴ*). As stated just above, however, these sound-*katakana* combinations were created for a limited cases of loanwords. As a matter of fact, the *kana*s (ァ(*a*), ィ (*i*), ェ (ɛ), ォ(ɔ)) reduced in size were used to indicate that *ɸ* is not immediately followed by [u] but exceptionally by other vowels. The use of フ presumably indicates that debuccalization did not apply even before those vowels and *ɸ* remains to be labial imitating the sound [f]. This view is supported by the fact that many old speakers use the conservative default pronunciation of フ as *ɸu* in foreign words like ***ɸu****aɴ* 'fan' (ファン) and ***ɸu****iɾumu* 'film' (フイルム), (Vance 1987:21) — note the use of unreduced ア (*a*) and イ (*i*). It therefore seems to be premature to draw any conclusion based upon the exceptional cases mentioned above.

5. Conclusion:

5.1. Summary:

The present study proposes an "underlying *ɸ* analysis" of the *ha-gyoo* consonants as an alternative to McCawley's (1968) "underlying *p* analysis." A series of production experiments yielded the following major findings. First, both lenition and fortition apply to all voiced obstruents in Japanese, leading them to be predominantly pronounced as continuants [z, ʒ, β̞, ð̞, ɣ̞] intervocalically but as non-continuants [dz, dʒ, b, d, g] post-nasally. Second, this positional contrast is not entirely categorical. Third, *z*/*dz* and *ʒ*/*dʒ* systematically showed fewer bursts than *b*/*β* , *d*/*ð*, and *g*/*ɣ* post-nasally as well as intervocalically, which suggests that the former group of voiced obstruents are realized as lenis/continuant more frequently than the latter irrespective of their positions. In the pursuit of explanations for these synchronic facts, various diachronic changes in the phonology and orthography of Japanese were investigated. A wealth of historical evidence, especially in relation to the "confusion of the four *kana*s," was found to eventually lend support to the proposed analysis of the *ha-gyoo* consonants.

While McCawley's "underlying *p* analysis" involves phonological changes within *ha-gyoo* in present-day Japanese that mimic their historical evolution, it was concluded that this analysis is ultimately out of sync with the historical facts. Recall that a major motivation for this analysis was the fact that *Rendaku* can be maintained as a straightforward voicing rule (i.e., one that does not alter place or manner of articulation). However, this motivation can also be achieved equally well, if not better, in the proposed analysis since the relevant sounds (*ɸ*, *β*, *b*) are all bilabial and occlusivization is added by an independent process of fortition. In addition, while morphophonemic alternations like *go-**ɸ**uɴ* 'five minutes' vs. *ip-**p**uɴ* 'one minute' were a synchronic empirical motivation for McCawley's analysis, the same alternations can also be captured by positing the phonological process *ɸ*→*p* to apply after a moraic consonant. Finally, the proposed analysis can rationalize the use of the *dakuten* and *han-dakuten* diacritics while the traditional analysis fails to do so. In other words, the proposed "underlying *ɸ* analysis" not only inherits the virtues of McCawley's "underlying *p* analysis" but also eliminates its inconsistencies.

5.2. Final remarks:

In closing, we would like to provide responses to two possible critiques of the approach pursued in this study. First, the present study's attempt to capture surface phonetic variability by appealing to morphophonology (both synchronically and diachronically) may be called into question. In particular, it may be objected that surface phonetic realizations cannot be reduced solely to abstract phonology, especially from the perspective of Pierrehumbert's (2016) claim that speakers accumulate episodic memory of phonetic details and this can affect their representations in memory. However, the arguments in this paper are not necessarily incompatible with this view. A key proposal of this paper is that the observed gradience and asymmetry in fortis-lenis alternations may have arisen when the positional conditions of the relevant phonological processes underwent diachronic change. These diachronic changes, in turn, may have been driven by speakers' episodic phonetic experiences unevenly affecting their knowledge of the phonological processes. This framing would be fully in line with Pierrehumbert's (2016) stance that the investigation of speech sounds should be pursued in a hybrid fashion, with

abstract phonological mechanisms placed at the center of inquiry and supplemented with the study of phonetic details and contextual features.

Second, in response to the proposal that the underlying form of *ha-gyoo* consonants is ϕ, it may be asked whether the *ha-gyoo* consonants need to have one underlying phoneme in the first place (as opposed to multiple). At a higher level, this claim may raise questions about the value and nature of the concept of underlying forms themselves. For example, what does it mean to postulate a particular underlying form? Is the primary deciding factor phonetic proximity, explanatory parsimony, typological generalizability, or phonology-orthography alignment? What if no single underlying form can be postulated that satisfies all of these conditions? These are all important questions that deserve further investigation, but they go beyond the scope of the current work. We hope that the proposal in this paper can stimulate more discussion and research on these theoretical issues.

6. Appendix A:

The stimuli used in Experiment #1 and Experiment #2 are listed below in (86) and (87), respectively. The column labels indicate which vowel (a/i/u/e/o) follows the target consonant. The row labels indicate the following three things:

Target consonant: ・*b*/*z*/*ʒ* in (86)
・*d*/*g* in (87)

Derivedness: ・r=*Rendaku*-voiced (i.e., in native and Sino-Japanese words)
・u=underlyingly voiced (i.e., in loanwords)

Position: ・N=post-nasal
・V=intervocalic
・C=citation form (i.e., target consonant in word-initial position) - (87) only

The contents of each cell contain the words used as stimuli in the experiment, written in normal Japanese orthography. The symbol | indicates the position immediately before the consonant in question.

(86) Stimuli for Experiment #1

		a	i	u	e	o
b	rN	郵便|箱 日本|晴れ 金|蝿	万|引き 観音|開き 記念|日	露天|風呂 堪忍|袋 本|降り	四人|部屋 甚|兵衛 トタン|塀	一番|星 初版|本 みりん|干し
	rV	筆|箱 五月|晴れ しょうじょう|蝿	福|引き 店|開き 土曜|日	檜|風呂 福|袋 大|降り	一人|部屋 吉|兵衛 土|塀	流れ|星 文庫|本 切り|干し
	uN	メン|バー ズン|バ ナン|バー	コロン|ビア キリン|ビール コン|ビーフ	ギャン|ブル アンサン|ブル モン|ブラン	ダン|ベル クラン|ベリー ボン|ベ	シン|ボル ジャン|ボ ドラゴン|ボール
	uV	キャ|バレー ハー|バー グロー|バル	ター|ビン 貸し|ビル ナイロ|ビ	アドリ|ブ サ|ブ タ|ブー	ノー|ベル賞 ブルー|ベリー ボケ|ベル	カー|ボン プレイ|ボーイ ベース|ボール
z	rN	寒|桜 小判|鮫 裁判沙|汰	河川|敷 どん|尻 豚|汁	ポン|酢 写真|好き 片面|刷り	みんみん|蝉 ガン|攻め 軍|勢	沿線|沿い どん|底 小紋|染め
	rV	八重|桜 甚平|鮫 表|沙汰	下|敷き 目|尻 鼻|汁	すし|酢 きれい|好き 二色|刷り	油|蝉 水|攻め 総|勢	海|沿い 舟|底 京|染め
	uN	カン|ザス タン|ザニア トラン|ザム	チンパン|ジー ベン|ジン パン|ジー	ジーン|ズ ブロン|ズ レン|ズ	ベン|ゼン シャン|ゼリゼ キン|ゼー	ゴルゴン|ゾーラ ベン|ゾール バトン|ゾーン
	uV	ユー|ザー ジグ|ザグ ブロポー|ザル	クレー|ジー ビ|ジター ブラ|ジル	バ|ズーカ レー|ズン ブラ|ズマ	ディー|ゼル イー|ゼル マドモア|ゼル	アマ|ゾン リ|ゾート ブル|ゾン
ʒ	rN	忍|者 銀|鮭 計算|尺	N/A	新|宿 臨|終 真|珠	N/A	避難|所 根|性 誕|生
	rV	王|者 塩|鮭 巻き|尺	N/A	原|宿 始|終 宝|珠	N/A	研究|所 素|性 養|生
	uN	ジン|ジャー ダウン|ジャケット ベン|ジャミン	チェン|ジ エン|ジン オレン|ジ	レモン|ジュース ホン|ジュラス 缶|ジュース	イン|ジェクター タン|ジェント エン|ジェル	エン|ジョイ パン|ジョー エルトン|ジョン
	uV	マネー|ジャー ティーンエー|ジャー メ|ジャーリーグ	チャー|ジ アベレー|ジ ダメー|ジ	りんご|ジュース スケ|ジュール モ|ジュール	エー|ジェント ネグリ|ジェ ターボ|ジェット	マ|ジョリティー ディスク|ジョッキー ダウ|ジョーンズ

(87) Stimuli for Experiment #2

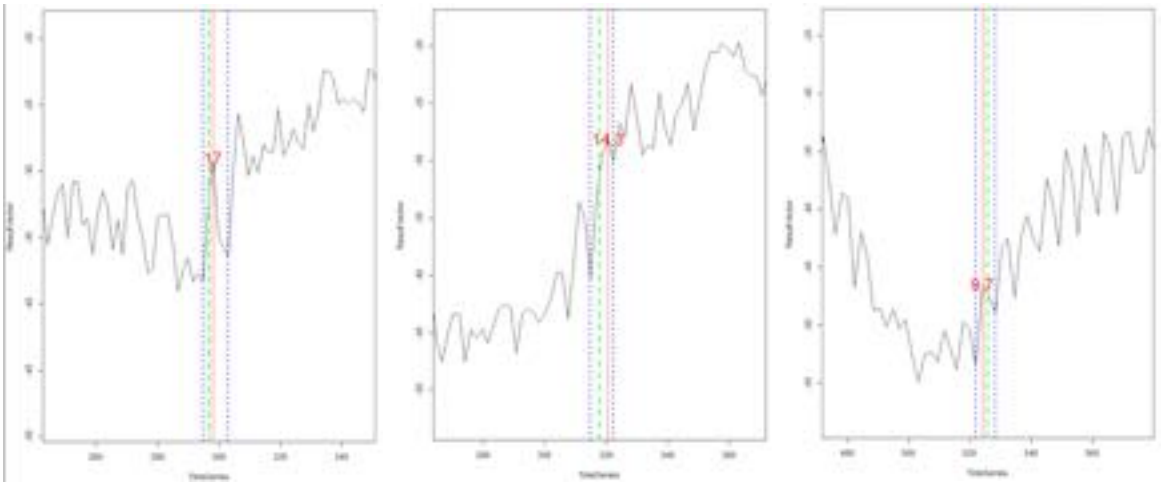
		a	i	u	e	o
d	rN	パン|種 シャボン|玉 陣|太鼓	N/A	N/A	禅|寺 彼岸|照り 弓|手	祇園|通り 御飯|時 寸|止め
	rV	物|種 ガラス|玉 大|太鼓	N/A	N/A	苔|寺 油|照り 熊|手	電車|通り 日暮れ|時 口|止め
	uC	ダンス ダム ダウン	N/A	N/A	デビュー デザイン デモ	ドア ドーム ドック
	uN	ラテン|ダンス 黒四|ダム トーン|ダウン	N/A	N/A	公園|デビュー 商品|デザイン 反戦|デモ	回転|ドア 溶岩|ドーム 人間|ドック
	uV	フォーク|ダンス 利根川|ダム コスト|ダウン	N/A	N/A	レコード|デビュー 宝石|デザイン 反米|デモ	木製|ドア 原爆|ドーム 心臓|ドック
g	rN	新|顔 番|傘 日本|髪	銀|狐 千|切り 人間|嫌い	玄関|口 番|組 天|草	寒|稽古 二段|跳り ロン|毛	他人|事 番|小屋 本|腰
	rV	古|顔 洋|傘 洗い|髪	古|狐 薄|切り 動物|嫌い	通用|口 骨|組み 餅|草	立ち|稽古 回し|跳り 赤|毛	隠し|事 山|小屋 逃げ|腰
	uC	ガス ガール ガード	ギター ギフト ギャング	グラス グッズ グループ	ゲーム ゲート ゲリラ	ゴリラ ゴルフ ゴール
	uN	メタン|ガス バトン|ガール スキン|ガード	ハワイアン|ギター お中元|ギフト イタリアン|ギャング	ワイン|グラス 受験|グッズ 支援|グループ	アクション|ゲーム 三番|ゲート 共産|ゲリラ	マウンテン|ゴリラ オープン|ゴルフ オウン|ゴール
	uV	排気|ガス カバー|ガール ボディー|ガード	フラメンコ|ギター ご婚礼|ギフト ストリート|ギャング	ビール|グラス 旅行|グッズ 武装|グループ	ソーシャル|ゲーム 中央|ゲート 革命|ゲリラ	ローランド|ゴリラ お座敷|ゴルフ ベスト|ゴール

7. Appendix B:

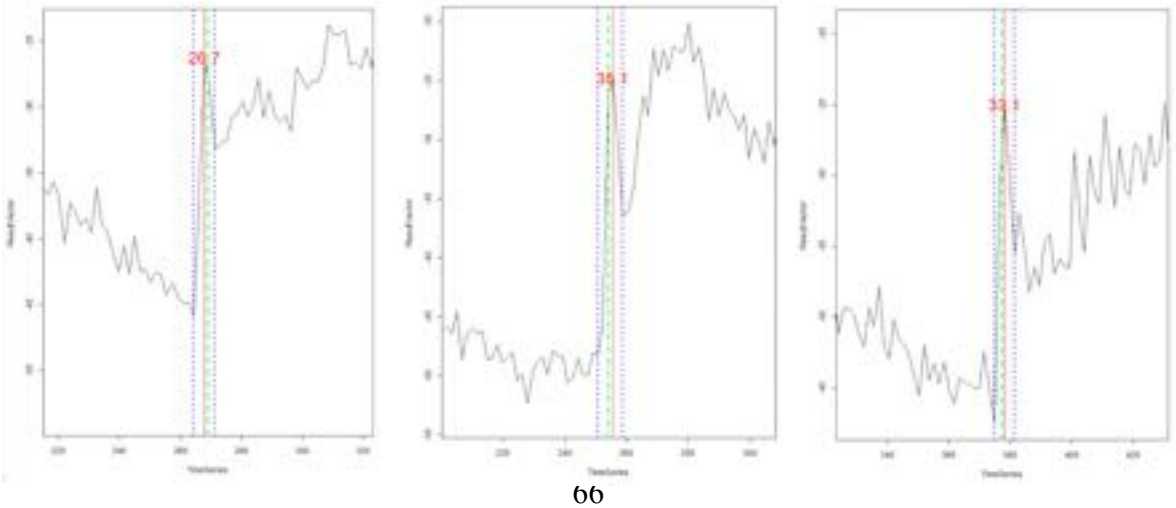
As described in Section 2.1.3, bursts were detected audiovisually by locating brief spurts of high-intensity energy in the spectrogram while listening to the corresponding soundfile. While the present study used a three-way contrast (yes/no/indeterminate) to minimize bias, the reliability of the resulting judgments may still be called into question. Thus, a post-hoc analysis was conducted to verify the extent to which the manual judgements agree with an alternative objective approach (implemented in R) based on automatic peak detection.

The plots in (88) and (89) below are a representative selection of recordings from Experiments 1 and 2, respectively, with each plot/recording coming from a different speaker. In each plot, the x axis corresponds to the time domain (zoomed in to the vicinity of the target sound), and the y axis represents the median amplitude (in db) across all frequencies at a given time point. The dashed (green) line corresponds to the position of the burst manually identifier by the human annotator. The two dotted (blue) lines indicate the valleys (i.e., local minima) to the left and right of this point. The solid (red) line indicates the peak (i.e., local maximum) between these two valleys. The value indicated in the middle of each plot is the sum of the amplitude rise (peak minus preceding valley) and amplitude fall (peak minus following valley), indicating the 'prominence' of the burst.

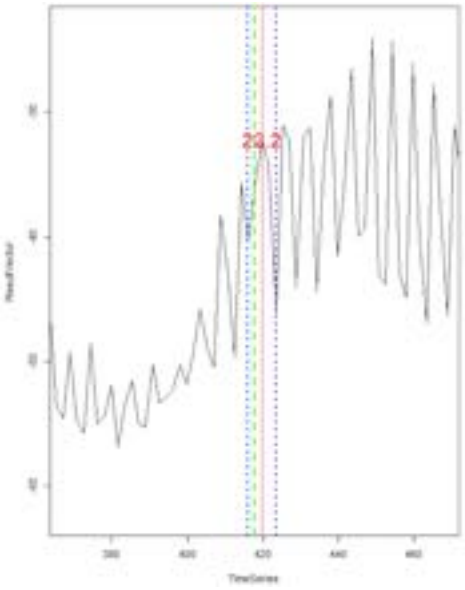
(88) a. yae_**z**_akura b. ʃio_**ʒ**_ake c. satsuki_**β**_aɾe



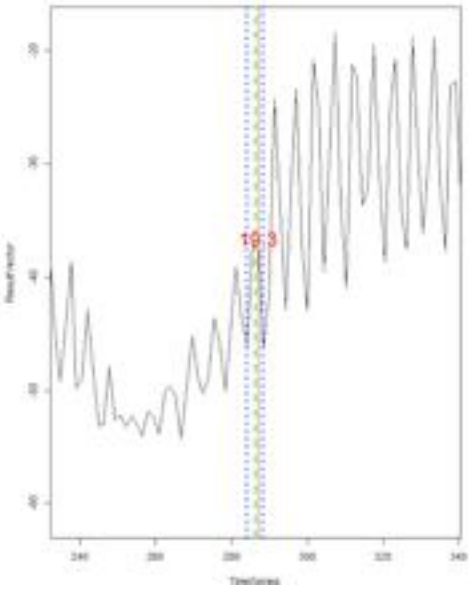
d. kaN_**dz**_akura e. giN_**dʒ**_ake f. nihoN_**b**_aɾe



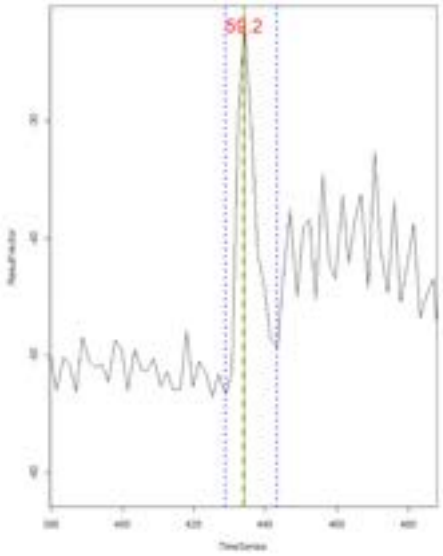
(89) a. gaɾasu_**ð**_ama



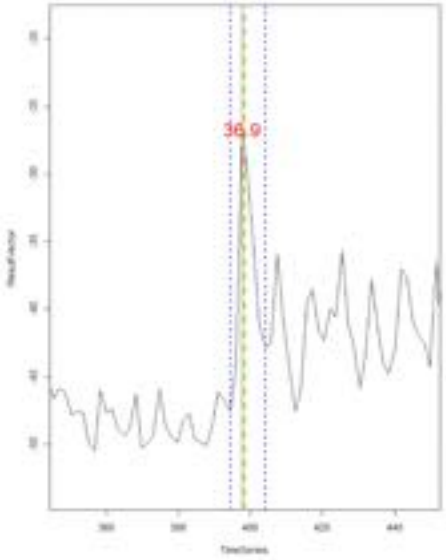
b. ɸuɾu_**ɣ**_ao



c. ʃaboN_**d**_ama



d. ʃiN_**g**_ao



Crucially, in every case, the manually-determined burst location is closely aligned to the automatically detected peaks. This corroborates the reliability of the analyses based on manual judgments for the present study. In this connection, it is worth noting that several other studies have also used manual judgements of burst presence, e.g. Maekawa (2010a), Maekawa (2010b) and Maekawa (2013) for Japanese and Kataoka (2010) for Northern Paiute..

Furthermore, note how, in each set of plots, the burst prominence (amplitude rise+fall) values are systematically larger for the second row (post-nasal) than the first row (intervocalic). This suggests that, even where bursts were detectable in intervocalic position, they were of weaker prominence compared to post-nasal position. While still tentative based on the post-hoc analysis of a handful of recordings, if valid, this conclusion would bolster the arguments made in the body of the paper regarding the presence vs absence of bursts.

References:

Amanuma, Yasushi, Kazuo Otsubo and Osamu Mizutani (1997) *Nihongo Onseigaku (Japanese Phonetics)*, Kuroshio Publishing, Tokyo.

Arisaka, Hideyo (1959) *On'inron: Zōhoban ["Phonology: Expanded Edition"]*, Sanseidō, Tokyo.

Asada, Kentaro (2000) "Nihongo-ni okeru Handakuonka-o meguru Mondai —Shoinshiryo-o Tegakari-toshite," *Kamakura Jidaigo Kenkyū ["Studies on the Language of the Kamakura Period"]*, 23, 816-835.

Crawford, Clifford (2009) *Adaptation and Transmission in Japanese Loanword Phonology*, Ph. D. dissertation, Cornell University.

Doi, Tadao (1955) *Nihon-Daibunten [日本大文典; Translation of Arte da Lingoa de Iapam by João Rodriguez]*, Sanseido, Tokyo.

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

Hamada, Atsushi (1954) "Hagyōon-no mae-no sokuon — p-on-no hassei ["Sokuon before ha-gyō sounds — The rise of the p-sound"]," *Kokugogaku ["Studies in the Japanese Language"]*, 16, 22-28.

Hamada, Atsushi (1960) "On'inshi," *Kokugo-to Kokubungaku ["Japanese Language and Literature"]*, 37.10, 17-28.

Hamada, Atsushi (1977) "Kokugoshi-no Shomondai — Yonjunen-no Sokatsu," *Kokugo-Kokubun*, 46.5, 577-596.

Hashimoto, Shinkichi (1928) "Ha-gyo Shiin-no Hensen-ni Tsuite [On the Transition of the Ha-column Consonants]," in Ichikawa, Sanki ed., *Okakura-sensei-Kinen-Ronbunshu*, Okakura-sensei Kanreki Shukugakai, Tokyo, 196-207.

Hashimoto, Shinkichi (1966) *Kokugo On'inshi (Kōgishū Ichi) ["The History of the Sounds of Japanese (Lecture Collection #1)"]*, Iwanami Shoten., Tokyo.

Hattori, Shirō (1951) *On'inron-to Seishohō ["Phonology and Orthography"]*, Kenkyūsha, Tokyo.

Hayata, Teruhiro (1977) "Seisei AKusento-ron [A Generatie Study of Accent]," in Ono, Susumu and Takeshi Shibata eds., *Nihongo 5: On'in*, Iwanami, Tokyo, 323-360.

Hizume, Shūji (2002) "Hagyoushiin-o meguru Yonshu-no 'yuuseika' ['On four types of vocalization in ha-gyou consonants']," *Ibarakidaigaku Jinbungakubu Kiyou*, 37, 97-118.

Hizume, Shūji (2003) "Seidaku bunka-to sokuon, hatsuon ["The establishment of sei-daku opposition and the phonetic grounds for sokuon and hatsuon development"]," *Kokugogaku ["Studies in the Japanese Language"]*, 54.2, 95-108.

Hoffmann, Johann Joseph (1868) *A Japanese Grammar*, E.J. Brill, Leiden.

Irwin, Mark (2001) *Loanwords in Japanese*, John Benjamins, Amsterdam.

Itô, Junko (1988) *Syllable Theory in Prosodic Phonology*, Garland Publishing, New York:.

Ito, Junko and Armin Mester (2003) *Japanese Morphophonemics*, The MIT Press, Cambridge MA.

Ito, Junko and Armin Mester (2008) "Lexical classes in phonology," in Miyagawa, Shigeru and Mamoru Saito eds., *The Oxford Handbook of Japanese Linguistics*, Oxford University Press, Oxford, 84–106.

Ito, Junko and Ralf-Armin Mester (1986) "The Phonology or Voicing in Japanese: Theoretical Consequences for Morphological Accessibility," *Linguistic Inquiry*, 17, 49-73.

Kamei, Takashi (1950) "Kenshukuryokoshu-o Chushinni Mita Yotsugana (The confusion of four kana characters based mainly upon Kenshukuryokoshu)," *Kokugogaku ["Studies in the Japanese Language"]*, 4, 75-88.

Kamei, Takashi (1956) "Gagyō-no kana ["Kana of the ga-gyō"]," *Kokugo-to Kokubungaku ["Japanese Language and Literature"]*, 33, 1-14.

Kamei, Takashi (1960) "Zaitoki-no 'Hongo Ha-ji-no On'-ni Kansuru Kaishaku [An Interpretation of "本郷波字音” in Zaitôki]," *Kokugogaku ["Studies in the Japanese Language"]*, 40, 126-131.
Kataoka, Reiko (2010) "Phonetic description of a three-way stop contrast in Northern Paiute," *University of California Berkeley Phonology Lab Annual Report*, 316-347.
Kawahara, Shigeto and Hideki Zamma (2016) "Generative treatments of rendaku and related issues," in Vance, Timothy J. and Mark Irwin eds., *Sequential Voicing in Japanese: Papers from the NINJAL Rendaku Project*, John Benjamins, Amsterdam, 13-34.
Kawakami, Shin (1977) *Nihongo Onsei Gaisetsu ["Survey of the Sounds of Japanese"]*, Ōfūsha, Tokyo.
Kingston, John (2008) "Lenition," *Selected proceedings of the 3rd conference on laboratory approaches to Spanish phonology, Cascadilla Proceedings Project*, 1–31,
Kiyose, Gisaburo N. (1985) "Evidence that Modern Japanese *h-* was *p-* in the Heian Period," *Onsei-no Kenkyu*, 21, 73-87.
Komatsu, Hideo (1981) *Nihongo-no Onin (Phonology in Japanese) — Nihongo-no Sekai (The World of Japanese) 7*, Chuo Koronsha, Tokyo.
Kubozono, Haruo (1999) *Nihongo-no Onsei (Speech Sounds in Japanese)*, Iwanami, Tokyo.
Labrune, Laurence (2012) *The Phonology of Japanese*, Oxford University Press, Oxford.
Mabuchi, Kazuo (1977) "Kokugo-no 'dakuon' ["Voiced obstruents in Japanese"]," in Kinenkai, Matsumura Akira Kyōju Kanreki ed., *Kokugogaku to kokugoshi : Matsumura Akira Kyōju kanreki kinen*, Meijishoin, Tokyo,
Maekawa, Kikuo (2010a) "Coarticulatory reinterpretation of allophonic variation: Corpus-based analysis of /z/ in spontaneous Japanese," *Journal of Phonetics*, 38, 360–374.
Maekawa, Kikuo (2010b) "Weakening of Stop Articulation in Japanese Voiced Plosives," *Onsee-kenkyuu [The Study of Sounds]*, 14.2, 1-15.
Maekawa, Kikuo (2013) "Notes on So-called Inter-speaker Difference in Spontaneous Speech: The Case of Japanese Voiced Obstruent," *INTERSPEECH 2013*, 3037-3041.
McCawley, James D. (1965) *The Accentual System of Standard Japanese*, Ph. D. dissertation, MIT.
McCawley, James D. (1968) *The Phonological Component of a Grammar of Japanese*, Mouton, The Hague.
Mitani, Shigeki, Toshihiro Kitama and Yu Sato (2006) "Voiceless affricate/fricative distinction by frication duration and amplitude rise slope," *The Journal of the Acoustical Society of America*, 120.3, 1600-1607.
Nakata, Norio (1972a) *Kōza kokugoshi: On'inshi, mojishi [A course on the history of Japanese: History of sounds, history of characters]*, Taishūkan Shoten, Tokyo.
Nakata, Norio (1972b) *Mojishi・Oninshi*, Taishukan, Tokyo Japan.
Nakata, Norio (1972c) "Sōsetsu [General remarks]," in Nakata, Norio ed., *Kōza kokugoshi: On'inshi, mojishi*, Taishūkan Shoten, Tokyo, 3-62.
Nihon Onsei Gakkai, ed. (1976) *Onseigaku Daijiten*, Sanshiisha, Tokyo.
Numoto, Katsuaki (1990) "Handakuonfu shijō-ni okeru tōin shiryō-no ichi ["The significance of the tō-in materials in the history of Japanese p-sound marks"]," *Kokugogaku ["Studies in the Japanese Language"]*, 162, 1-12.
Numoto, Katsuaki (1992) "Nihongo-no Hyoki," in Nagao, Shoso ed., *Nihongogaku*, Izumi Shoin, Tokyo, 65-102.
Numoto, Katsuaki (1993) "Dakuon-jibo-kara Dakushouten-e: Dakuten-no kigenron-zokuchou ["On the Origin of Dakuten"]," *Kokugogaku ["Studies in the Japanese Language"]*, 172, 15-28.
Ohno, Kazutoshi (2000) "The Lexical Nature of *Rendaku* in Japanese," in Nakayama, Mineharu and Charles J. Quinn Jr. eds., *Japanese/Korean Linguistics 9*, CSLI, Stanford, 151-164.

Okumura, Mitsuo (1972) "Kodai-no on'in [Sounds of the ancient period]," in Nakata, Norio ed., *Kōza kokugoshi: On'inshi, mojishi*, Taishūkan Shoten, Tokyo, 63-172.
Rosen, Eric (2001) *Phonological processes interacting with the lexicon: Variable and non-regular effects in Japanese Phonology*, Ph. D. dissertation, University of British Columbia.
Sato, Kiyoji (ed.) (1977) *Kokugogaku Kenkyu Jiten (An Encyclopedia of Japanese Language Studies)*, Meiji Shoin, Tokyo.
Takayama, Michiaki (1992-3) "Rendaku-to Renjyoodaku," *Kuntengo to kuntenshiryoo*, 88, 115-124.
Takayama, Tomoaki (1994) "Fukugougo-niokeru Sokuon-no Sounyuu-nituite — Mou Hitotsu-no Rendaku ["On the gemination in compounds — Another case of Rendaku"]," in Iinkai, Morino Muneaki Kyouju Taikan Kinen Ronshu Henshu ed., *Gengo, Bungaku, Kokugokyouiku: Morino Muneaki Kyouju Taikan Kinen Ronshu*, Sanseido, Tokyo, 347-360.
Takayama, Tomoaki (2003) "Kenshukuryōkoshū kara ukagaeru zenbion'yōso no ichi kyokumen [An aspect of the de-prenasalization seen in Kenshukuryōkoshū]," *Kagawa Daigaku Kokubun Kenkyū*, 28, 94-101.
Takayama, Tomoaki (2015) "Historical Phonology," in Kubozono, Haruo ed., *Handbook of Japanese Phonetics and Phonology*, Mouton De Gruyter, Berlin, 402-417 (Electronic version).
Toyama, Eiji (1972) "Kindai-no on'in [Sounds of the modern period]," in Nakata, Norio ed., *Kōza kokugoshi: On'inshi, mojishi*, Taishūkan Shoten, Tokyo, 173-268.
Tsukishima, Hiroshi (1964) *Kokugogaku (The Study of the Japanese Language)*, Tokyo Daigaku Shuppankai, Tokyo.
Ueda, Kazutoshi (1903) "P-on-ko [On p-sound]," *Kokugo-no-tame*, 2, 32-39.
Vance, Timothy (2008) *The Sounds of Japanese*, Cambridge University Press, Cambridge.
Vance, Timothy J. (1987) *An Introduction to Japanese Phonology*, State University of New York Press, Albany.
Vance, Timothy J. (2014) "If rendaku isn't a rule, what in the world is it?," in Kabata, Kaori and Tsuyoshi Ono eds., *Usage-based Approaches to Japanese Grammar: Towards the understanding of human language*, John Benjamins, Amsterdam/Philadelphia, 137-151.
Vance, Timothy J. and Mark Irwin (2016) *Sequential Voicing in Japanese: Papers from the NINJAL Rendaku Project*, John Benjamins, Amsterdam.
Wang, Li (2012) *Hanyu-yuyin-shi, Beijing First Edition*, Zhong Hua Shuju, Beijing, China.
Yamada, Takao (1904a) "Rendakuon-no Hassei ["The Rise of the Sounds Voiced by Renadaku"]," *Kokugakuinzasshi*, 10.8, 33-43.
Yamada, Yoshio (1904b) "Rendaku-ni Tsuite [On the Rise of Rendaku-on]," *Kokugakuinn-Zasshi*, 10.8,
Yamaguchi, Toshiko (2007) *Japanese linguistics: An introduction*, Continuum, London.