Shigeto Kawahara

# 1 The phonetics of *sokuon*, or geminate obstruents

## 1 Introduction

Japanese has a phonemic contrast between short and long nasal and obstruent consonant series, as exemplified by minimal pairs like [ka**t**a] 'frame' vs. [ka**tt**a] 'bought' and [ha**t**o] 'dove' vs. [ha**tt**o] 'hat'.[1] Short consonants are generally called "singletons", whereas long consonants are called "geminates", geminate obstruents, or obstruent geminates (see also Kawagoe, this volume). In the traditional literature on Japanese phonetics and phonology, the first half of obstruent geminates is called "*sokuon*" for which the symbol /Q/ is often used; in the Japanese orthographic system, this coda part is represented by "small tsu". Nasal geminates or their coda portions are called *hatsuon*; in the traditional literature they are represented by /N/. This chapter focuses on obstruent geminates. Henceforth, the term "geminate" refers specifically to obstruent geminates or *sokuon*, unless otherwise noted. This chapter provides an overview of the acoustic, perceptual, and articulatory characteristics of Japanese geminates.[2]

---

**1** There is no phonemic contrast between short and long approximants (liquids and glides) in Japanese (see Kawagoe, this volume). Geminate approximants can however occur in emphatic forms (e.g. [kowwai] 'very scary' is an emphatic form of [kowai] 'scary'). See Aizawa (1985), Kawahara (2001), Kawahara and Braver (2014), and section 5.3 for the non-structure preserving nature of this emphatic gemination in Japanese. For the phonetic reasons that may possibly underlie the prohibition against lexical approximant geminates, see Kawahara, Pangilinan, and Garvey (2011), Kawahara (2012), Podesva (2000) and Solé (2002).

**2** Primarily due to limitation of the author's expertise, L2 learning of Japanese geminates is not covered in this paper. Readers are directed to the following references: Han (1992); Motohashi-Saigo and Hardison (2009); Oba, Brown, and Handke (2009); Tajima et al. (2008), several papers in a special issue of *Onsei Kenkyū* 11:1 (Kubozono 2007), those cited therein, as well as Hirata (this volume). Another topic that this chapter does not cover is a gemination pattern found in the process of loanword adaptation (e.g. [bakkɯ] 'back' < English *back*), which arguably has a perceptual basis (e.g. Kawagoe and Takemura 2013; Takagi and Mann 1994, though cf. Kubozono, Ito and Mester 2008). See Kawagoe (this volume) and Kubozono (this volume) for further discussion on this phenomenon.

This chapter does not deal with long vowels, although many issues discussed for geminate consonants in this paper are also relevant to long vowels. Here I list some key references. For general durational properties of long vowels in Japanese, see Braver and Kawahara (2014); Han (1962); Hoequist (1982); Kawahara and Braver (2013); Mori (2002) and Port, Dalby, and O'Dell (1987); for the effect of speech rate on long vowel production and perception, see Hirata (2004) and Hirata and Lambacher (2004); for secondary, non-durational acoustic correlates and their perceptual impacts, see Behne et al. (1999); Hirata and Tsukada (2009) and Kinoshita, Behne, and Arai (2002).

The structure of this paper is as follows. Section 2 discusses acoustic correlates of a singleton/geminate contrast in Japanese. The primary acoustic correlate exploited by Japanese speakers is constriction duration; other acoustic correlates include various durational correlates (e.g. duration of preceding vowel) and non-durational correlates (e.g. spectral properties in surrounding vowels). Section 2 also discusses other topics including the search for invariance and manner effects, as well as comparison of Japanese with other languages. Section 3 provides an overview of the experiments on the perception of geminates in Japanese. It discusses the effect of constriction duration as the primary perceptual cue, and also discusses how the duration of surrounding intervals affects the perception of geminates. Section 4 provides an overview of the literature on the articulation of Japanese geminates. Several issues that require further investigation are identified throughout the paper, and Section 5 raises several other issues that are not covered in the rest of the paper.

# 2 The acoustic characteristics of geminates in Japanese

## 2.1 The primary acoustic correlate: constriction duration

Japanese is often assumed to be a mora-timed language (see Warner and Arai 1999 for a review; see also Otake, this volume, on mora-timing); geminates are moraic, while singletons are not; for example, disyllabic words containing a geminate like [katta] 'bought' or [hatto] 'hat' have three moras. Reflecting their moraic nature, geminate consonants in Japanese have a longer consonantal constriction. Acoustically, the primary correlate of a singleton-geminate contrast is a difference in constriction duration – i.e. for stops, it is closure duration and for fricatives, it is frication duration. (In this paper, "duration" refers phonetic measures and "length" refers phonological contrast; "constriction" refers to both stop closure and narrow aperture for fricatives).[3]

Before proceeding to the discussion, there is one remark about what is meant by a particular acoustic correlate being "primary". The concept of being "primary" can mean several different things. A primary acoustic correlate can be used to mean an acoustic parameter that is invariant across speakers, speech styles, phonological contexts, or even across languages; a "primary" cue is also used to mean that it constitutes the most important perceptual cue for listeners, one that dominates other secondary cues (Lahiri and Hankamer 1988) so that secondary cues are only exploited when the target stimuli are ambiguous in terms of the primary cue, distributing around a range that is not found in natural speech (Hankamer, Lahiri, and

3 For affricates, the primary acoustic correlate seems to lie in the difference in the closure duration, and not in frication duration (Oba, Brown, and Handke 2009). See section 2.3.2.

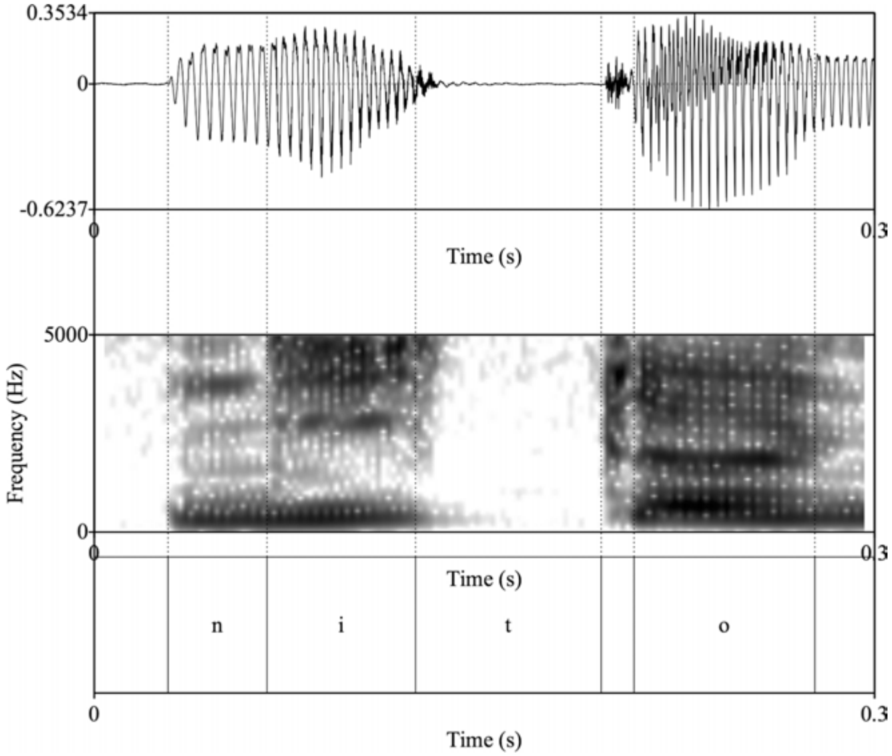


**Figure 1:** A singleton [t] in Japanese. Produced by a female native speaker of Japanese. The time scale is 300ms

Koreman 1989; Picket, Blumstein, and Burton 1999). For a general discussion on primacy of cues, see Abramson and Lisker (1985); Stevens and Blumstein (1981); Stevens and Keyser (1989); Whalen et al. (1993) and others; for a discussion of primacy in the context of length distinctions, see Abramson (1992); Hankamer, Lahiri, and Koreman (1989); Idemaru and Guion (2008); Lahiri and Hankamer (1988); Picket, Blumstein, and Burton (1999) and Ridouane (2010). Ridouane (2010) argues that cross-linguistically, differences in constriction duration are the most consistent acoustic correlates of singleton-geminate contrasts.

With this said, the primary acoustic correlate of Japanese geminates is greater duration compared to singletons: geminate consonants are characteristically longer than singleton consonants. Figures 1 and 2 show illustrative waveforms and spectrograms of a singleton [t] and a geminate [tt] in Japanese (with the same time scale of 300ms). As we can see, the geminate [tt] has a longer closure than the singleton [t].

Many acoustic studies have investigated the durational properties of singleton-geminate contrasts in Japanese, and Table 1 summarizes their findings. This summary shows that geminate stops are generally at least twice as long as corresponding singleton stops, and can sometimes be as three times as long, regardless of the

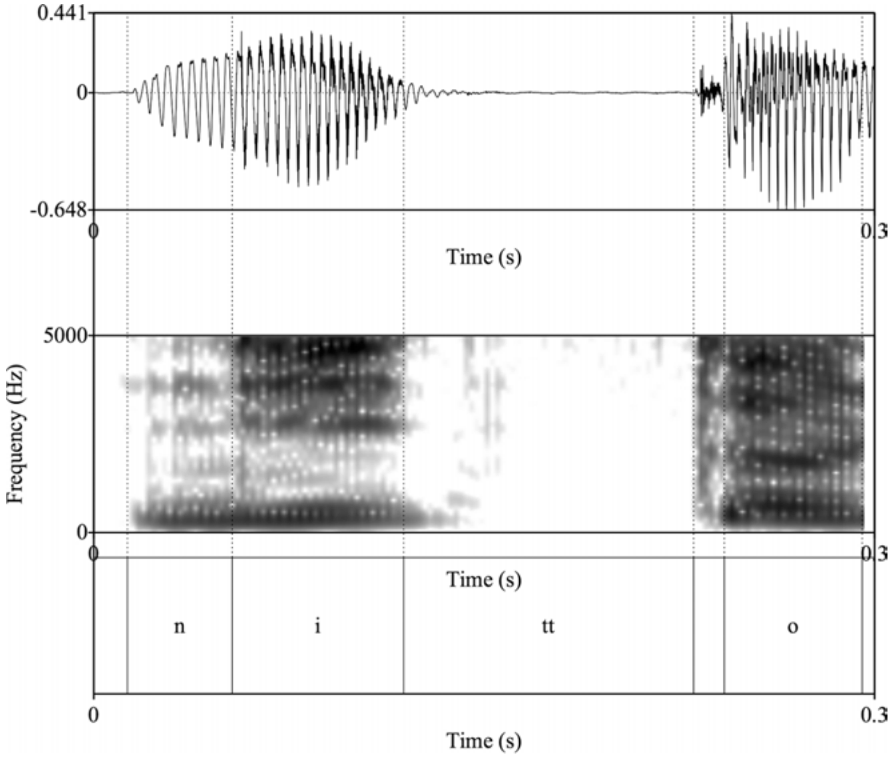


**Figure 2:** A geminate [tt]. The time scale is 300ms

place of articulation or voicing status of the consonants (though see section 2.3 for further discussion on the manner effect on geminate duration).

## 2.2 Secondary acoustic correlates

As with many other phonological contrasts, a singleton-geminate contrast is acoustically manifested not only by constriction duration, but by multiple other acoustic properties as well. (Multiplicity of acoustic correlates for phonological contrasts has been an important topic throughout the history of the phonetic theory; see, for example, Abramson 1998; Kingston and Diehl 1994; Lisker 1986 and references cited therein.)

### 2.2.1 Other durational correlates

In Japanese, vowels are longer before geminates than before singletons (Campbell 1999; Fukui 1978; Han 1994; Hirata 2007; Hirose and Ashby 2007; Idemaru and Guion

**Table 1:** Summary of the previous studies on closure duration of singleton and geminate stops and their ratios in Japanese. Duration measures are in miliseconds. SD = standard deviation; MoE = margin of error for 95% confidence intervals. Sing = singleton; Gem = geminate; VOT = Voice Onset Time; vls = voiceless; vcd = voiced

| Sources | Sing duration | Gem duration | Ratio | Note |
|---|---|---|---|---|
| Han (1962) | – | – | 2.6–3.0 | based on small N |
| Homma (1981) | [p]: 77 | [pp]: 183 | 2.38 | 4 speakers |
| | [b]: 55 | [bb]: 159 | 2.89 | |
| | [t]: 62 | [tt]: 170 | 2.74 | |
| | [d]: 35 | [dd]: 144 | 4.11 | |
| | [k]: 61 | [kk]: 175 | 2.87 | |
| | [g]: 41 | [gg]: 134 | 3.27 | |
| Beckman (1982) | | | | (SD), 5 speakers |
| | [k]: 89 (17) | [kk]: 195 (32) | 2.25 | VOT included |
| | [k]: 64 (15) | [kk]: 171 (32) | 2.79 | VOT excluded |
| Port et al. (1987) | | | | (SD), 10 speakers |
| | [k]: 65 (12) | [kk]: 149 (25) | 2.29 | ɯ_ ɯ |
| | [k]: 66 (14) | [kk]: 146 (28) | 2.21 | a_ ɯ |
| Han (1994) | | | | (SD), 10 speakers |
| (see also Han 1992) | [p]: 76.3 (5.6) | [pp]: 195.9 (21.9) | 2.57 | sɯ_ai |
| | [p]: 72.9 (9.7) | [pp]: 205.4 (29.9) | 2.82 | sɯ_ori |
| | [t]: 71.5 (7.4) | [tt]: 192.3 (27.2) | 2.69 | i_e |
| | [t]: 53.5 (8.0) | [tt]: 166.6 (24.1) | 3.11 | ki_e |
| | [t]: 57.9 (10.2) | [tt]: 174.5 (21.5) | 3.01 | ʃi_ei |
| | [t]: 52.7 (8.0) | [tt]: 170.9 (25.8) | 3.24 | ki_e |
| | [t]: 68.2 (9.0) | [tt]: 189.8 (28.5) | 2.78 | i_a |
| | [k]: 63.5 (8.5) | [tt]: 178.2 (22.5) | 2.81 | yo_a |
| | [k]: 57.5 (8.5) | [tt]: 175.8 (30.9) | 3.06 | ʃi_e |
| | [k]: 79.4 (6.6) | [kk]: 198.7 (24.6) | 2.50 | ha_eN |
| Kawahara (2006a) | vls: 59.9 (2.1) | vls: 128.6 (3.1) | 2.15 | (MoE), 3 speakers |
| | vcd: 42.3 (1.7) | vcd: 113.1 (3.0) | 2.67 | |
| Hirose and Ashby (2007) | vls: 60.5 | vls: 114.2 | 1.89 | 3 speakers |
| | vcd: 44 | vcd: 108 | 2.45 | |
| Idemaru & Guion (2008) | 69 (28) | 206 (45) | 2.99 | (SD), 6 speakers all stop consonants |

2008; Kawahara 2006a, 2013b; Kawahara and Braver 2014; Ofuka 2003; Port, Dalby, and O'Dell 1987; Takeyasu 2012).[4] Port, Dalby, and O'Dell (1987) found, for example, that [ɯ] is on average 68ms before singleton [k] and 86ms before geminate [kk]; i.e. that [ɯ] is 18ms longer on average before geminates. Kawahara (2006a) found

4 Vowels are also longer in closed syllables before a so-called moraic nasal (or *hatsuon*) – i.e. in (C)VN – than in open syllables – i.e. in (C)V (Campbell 1999). This observation indicates that this lengthening is due to a general, syllable-based phenomenon. The pre-geminate lengthening can also block otherwise productive high vowel devoicing between two voiceless consonants (Han 1994; Takeyasu 2012; see also Fujimoto, this volume).

similarly that vowels before voiceless singletons are on average 36.9ms while those before voiceless geminates are 53.4ms. Furthermore, some studies even found that in $C_1VC_2V$ contexts, $C_1$ is longer when $C_2$ is a geminate than when $C_2$ is a singleton (Han 1994; Port, Dalby, and O'Dell 1987) (cf. Takeyasu 2012 who found the opposite, shortening pattern; Hindi shows the same lengthening pattern: Ohala 2007).

On the other hand, vowels that follow geminate/singletons show the reverse pattern: those that follow geminate consonants are shorter than those that follow singleton consonants (Campbell 1999; Han 1994; Hirata 2007; Idemaru and Guion 2008; Ofuka 2003). Han (1994) found the shortening of post-geminate vowels (and sometimes also the following word-final moraic nasals) by 9ms. In an acoustic study reported in Idemaru and Guion (2008), the mean duration of the following vowel is 63ms after geminates and 76ms after singletons. As explicitly noted by Hirata (2007), however, this difference in duration of the following vowels is less substantial and less consistent than the difference in the preceding vowel.

Finally, one may expect that Voice Onset Time (VOT) – an interval between the release of the closure and the onset of voicing of the following vowel – would be longer for geminate stops than for singleton stops, because longer closure would result in higher pressure build-up behind the stop occlusion. However, this expectation does not seem to hold: in Han (1994), VOT is slightly shorter for geminates than for singletons; in other studies (Hirata and Whiton 2005; Homma 1981), the relationship is inconsistent. See Kokuritsu Kokugo Kenkyūjo (1990) for the data on the intraoral air pressure rise in Japanese consonants, which indeed shows that geminates do not involve higher intraoral air pressure rise.

### 2.2.2 Other non-durational, acoustic correlates

Several studies have investigated other non-durational, acoustic correlates of a singleton-geminate contrast in Japanese. Their findings are summarized in Table 2.

As observed in Table 2, Japanese geminates are associated with various non-durational cues. Given that, in addition to the primary acoustic correlate of constriction duration, there are a number of acoustic cues that are associated with Japanese geminates, they cannot be merely characterized as "long consonants".

A remaining question therefore is how to represent Japanese geminates phonologically. Many possibilities exist in answer to this question, such as (i) double consonants (often assumed in phonemic representation/transcription), (ii) moraic consonants (Hayes 1989), (iii) a special /Q/ phoneme – or *sokuon* – as assumed in the traditional literature (e.g. Hattori 1984), or (iv) a special syllable concatenater (Fujimura and Williams 2008). This issue should continue to be discussed in relation to the phonological behavior of Japanese geminates (see Kawagoe, this volume), as well as to the theory of phonetic implementation of phonological representations.

**Table 2:** A summary of other, non-durational, acoustic correlates of Japanese geminates. Reference keys: F = Fukui (1978), I&G = Idemaru & Guion (2008), O = Ofuka (2003), K = Kawahara (2006a). See the original papers for the details of the measurement procedures

| | Patterns | References |
|---|---|---|
| Intensity | – The mean intensity difference between the surrounding vowels is larger across geminates. | I&G, O |
| F0 | – F0 drop (a correlate of a lexical accent – see Kawahara, this volume) is larger across geminates. | I&G, O, K |
| | – F0 falls toward geminates in unaccented disyllabic words. | F |
| F1 | – F1 is lower after geminates. | K |
| Spectral tilt | – H1-A1 is smaller for vowels after geminates (i.e. vowels are creakier). | I&G |

### 2.2.3 The search for invariance

One general research program in phonetics is the search for invariance (Stevens and Blumstein 1981). The issue addressed in this program is whether, for each phonological distinction, there exists any acoustic correlate that is invariant across phonological contexts, individual speakers, and speech styles, etc., and if so, what those invariant acoustical properties are. This issue is particularly important for a singleton-geminate contrast, because, although geminates are longer than singletons given the same speech rate, geminates in fast speech styles can be shorter than singletons in slow speech styles (Hirata and Whiton 2005; Idemaru and Guion-Anderson 2010).[5]

Usually proposals for invariant measures take the form of a relationship between more than one acoustic parameter. The general idea behind these studies on phonological contrasts based on durations is rate normalization – listeners normalize the duration of incoming acoustic signal according to the speech rate, which can be (unconsciously) inferred from the duration of other intervals (Miller and Liberman 1979; Pickett and Decker 1960). For example, when a preceding vowel sounds short, a listener may perceive that the speaker is speaking fast, and as a result even a phonetically short interval may be interpreted as phonologically long.[6]

5 It has been observed in other languages (Italian: Pickett, Blumstein, and Burton 1999 and Persian: Hansen 2004) that geminates are more susceptible to change in duration due to speech rate than singletons are. This asymmetry seems to hold in the Japanese data as well (Hirata and Whiton 2005; Idemaru and Guion-Anderson 2010).

6 An alternative theory is auditory durational contrast. This auditory mechanism (more or less automatically) renders an interval to sound longer next to a shorter interval (this mechanism is sometimes referred to as "durational contrast"). This mechanism is arguably not specific to speech, as it applies to the perception of non-speech stimuli (Diehl and Walsh 1989; Kluender, Diehl, and Wright 1988). It is beyond the scope of this paper to compare these two theories (for further discussion on this debate, see Diehl, Walsh, and Kluender 1991; Fowler 1990, 1991, 1992; Kingston et al. 2009).

Several relational acoustic measures have been proposed as an invariant measure that distinguishes singletons from geminates across different speech rates. Hirata and Whiton (2005) recorded various disyllabic tokens of singletons and geminates in nonce words and real words in three speech styles (slow, normal, and fast), and considered three measures: (i) raw closure duration, (ii) $C/V_1$ ratio (the ratio between the target consonant and the preceding vowel), and (iii) C/W(ord) ratio. Hirata (2007) and Hirata and Forbes (2007) followed up on this study and considered three more measures: (i) $C/V_2$ ratio ($V_2$ = the following vowel), (ii) V-to-V interval (i.e. added durations of preceding vowel, constriction and VOT)[7] and (iii) VMora (V-to-V interval divided by average mora duration). Idemaru and Guion-Anderson (2010) tested yet a few more relational measures: $C/V_1$, $C/C_1V_1$, $C/V_2$, and $C/(C + V_2)$ (where C is the target consonant, $C_1$ and $V_1$ are the preceding consonant and vowel, and $V_2$ is the following vowel), in addition to those already tested by Hirata and Whiton (2005) (specifically, raw closure duration and C/W ratio). After recording their own various tokens of singletons and geminates in three speaking rates, for each measure, Idemaru and Guion-Anderson (2010) tested classification accuracy percentages based on raw values as well as *z*-transformed (normalized) values within each speaker. Finally, in the most recent study on this topic, Hirata and Amano (2012) introduced a yet new notion, subword, which is a disyllabic (C)V (C)CV sequence, which includes the target singleton and geminate consonant medially. This notion is equivalent to C/W in Hirata and Whiton's (2005) work, as they used only disyllabic words.

All of these studies used discriminant analyses for each proposed measure to calculate how many percentages of tokens are accurately classified as a member of the intended category. The classification accuracy percentages of all the measures in these studies are summarized in Table 3.

One tendency that is clear from Table 3 is that relational measures generally classify singletons from geminates better than raw durational values do. Just which relational measure best cross-classifies Japanese singletons from geminates is an interesting topic for on-going and future research. We cannot also deny the possibility that there are other measures, relational or not, which better cross-classify Japanese singletons and geminates, which are yet to be uncovered.[8]

Another important issue is the perceptual relevance – or reality – of the relational, invariant acoustic measures: whether Japanese listeners exploit relational, acoustic measures, and if so, which measures are they sensitive to. For example,

---

**7** For example, given [kata], the V-to-V interval is [at], and given [katta], the V-to-V interval is [att].

**8** Other relational invariant measures proposed for length contrasts in other languages include $C/V_1$ ratio for Italian (Pickett, Blumstein, and Burton 1999), vowel to rhyme duration ratio for Icelandic (Pind 1986) (in which long vowels and geminates are more or less in a complementary distribution), and the ratio of the closure duration to the syllable duration in Persian (with some further complications) (Hansen 2004).

**Table 3:** A summary of classification accuracy percentages in the five studies cited in the text (chronologically ordered). See text for explanations of each measure

| | |
|---|---|
| Hirata and Whiton (2005) | |
| raw C duration | 82.2% (nonce words) and 81.4% (real words) |
| $C/V_1$ ratio | 92.1% (nonce words) and 91.3% (real words) |
| C/W | 98% (nonce words) and 95.7% (real words) |
| Hirata (2007) | |
| $C/V_2$ ratio: | 98.9% (nonce words) and 98.8-98.9% (real words) |
| Hirata and Forbes (2007) | |
| V-to-V interval | 75.5% |
| VMora | 99.6% |
| Idemaru and Guion-Anderson (2010) | |
| $C/V_2$ | 83.7% (raw) and 85.5% (normalized) |
| $C/C_1V_1$ (mora) | 92.6% (raw) and 94.5% (normalized) |
| $C/V_2$ | 94.1% (raw) and 94.9% (normalized) |
| $C/(C+V_2)$ | 92.3% (raw) and 93.0% (normalized) |
| C/Word | 96.3% (raw) and 96.8% (normalized) |
| raw C duration | 87.2% (raw) and 88.3% (normalized) |
| Hirata and Amano (2012) | |
| C/W | 97.5% (nonce words) and 93.9% (real words) |
| C/Subword (Subword=CV(C)CV) | 97.6% (nonce words) and 93.6% (real words) |

Idemaru and Guion-Anderson (2010) followed up their acoustic study with a perception test, which showed that while preceding mora ($C_1V_1$) duration significantly affects the perception of geminacy, whereas the following materials ($C/V_2$ ratio) do so only marginally, despite the fact that ratios involving these two factors yielded comparable accuracy percentages in production (see Table 3). See also Amano and Hirata (2010) and Otaki (2011) and section 3.2 for further discussion on the relationship between production and perception, especially in terms of contextual effects on the perception of length contrasts.

## 2.3 Manner and voicing effects

One issue that has received relatively less attention in the previous literature on Japanese geminates is the comparison of different manners of geminates in Japanese. Most previous acoustic studies on Japanese have investigated oral stops (Beckman 1982; Han 1992, 1994; Hirata and Whiton 2005; Hirose and Ashby 2007; Homma 1981; Idemaru and Guion 2008; Kawahara 2006a), although some studies included geminates of various manner types (e.g. Han 1962 measured oral stops and nasals; Campbell 1999 measured stops and some fricatives). Other languages that have been

**Table 4:** The effects of manner of articulation on the duration of singletons and geminates in Japanese (margin of error for 95% confidence intervals.)

| Segment | Singleton | Geminate | Ratio |
|---|---|---|---|
| [p] | 77.3 (7.8) | 129.6 (8.1) | 1.68 |
| [t] | 55.5 (4.6) | 124.4 (7.3) | 2.24 |
| [k] | 67.3 (7.1) | 128.7 (7.1) | 1.91 |
| [b] | 53.1 (3.8) | 131.4 (8.8) | 2.47 |
| [d] | 36.6 (1.9) | 116.0 (10.4) | 3.16 |
| [g] | 52.1 (3.7) | 115.0 (13.2) | 2.20 |
| [ɸ] | 83.5 (4.8) | 144.7 (7.4) | 1.73 |
| [s] | 83.2 (4.6) | 134.5 (7.0) | 1.62 |
| [ʃ] | 85.9 (5.7) | 138.4 (7.3) | 1.61 |
| [ç] | 63.4 (2.5) | 132.0 (6.2) | 2.08 |
| [h] | 72.2 (4.2) | 143.7 (6.4) | 1.99 |

studied in this light – manner effects on geminate contrasts – include Italian (affricates: Faluschi and Di Benedetto 2001; fricatives: Giovanardi and Di Benedetto 1998; nasals: Mattei and Di Benedetto 2000; see also Payne 2005), Cypriot Greek (Tserdanelis and Arvaniti 2001), Guinaang Bontok (Aoyama and Reid 2006), Finnish (Lehtonen 1970), Buginese, Madurese, and Toba Batak (Cohn, Ham, and Podesva 1999).

### 2.3.1 Fricative geminates

Japanese allows both (voiceless) stops and fricatives to contrast in geminacy. As in other languages (Lehiste 1970), singleton fricatives are generally longer than singleton stops in Japanese (Beckman 1982; Campbell 1999; Port, Dalby, and O'Dell 1987; Sagisaka and Tohkura 1984). As a result, geminate/singleton duration ratios are smaller for fricatives than for stops. Table 4 reports unpublished data collected by the author based on three female Japanese native speakers. All speakers were in their twenties at the time of recording, and the recording took place in a sound-attenuated room. Each target sound was pronounced in a nonce word frame [ni_o] (for most cases), itself being embedded in a frame sentence. Accents were always placed on the initial syllables. All three speakers repeated 10 repetitions of all tokens.[9]

Table 4 shows the results of duration measurements (for stops, VOT's were not included in the closure duration, as in many studies cited in Table 1). Duration ratios

9 I am grateful to Kelly Garvey and Melanie Pangilinan for their help with this acoustic analysis. This project also measured the duration of singleton and geminate nasals. The result shows that the geminate/singleton duration ratio for [n] was about 2.2 (Kawahara 2013a).

are highest for voiced stops than for voiceless stops (see also Homma 1981 and Hirose and Ashby 2007 for the same finding), which are also generally higher than for fricatives (except for [ç] and [h]).[10]

One phonological importance of the difference between stop pairs and fricative pairs is that the length contrast may be less perceptible for fricatives than for stops.[11] This less perceptible contrast of fricative pairs may lead to a diachronic neutralization (Blevins 2004) and/or avoidance of fricative geminates in synchronic phonological patterns (Kawahara 2006b, 2013b) due to a principle of contrastive dispersion to avoid contrasts that are not very well perceptible (Engstrand and Krull 1994; Flemming 2004; Liljencrants and Lindblom 1972; Lindblom 1986 and references cited therein; see also Martin and Peperkamp 2011 for a recent review on the effect of speech perception on phonological patterns.).

### 2.3.2 Affricate geminates

Affricates ([ts]) are not contrastive in the native phonology of Japanese, appearing as an allophonic variant of /t/ before [ɯ] (see Pintér, this volume); Geminate [ts], however, appears marginally in some borrowing as in [kjattsɯ] "cats" (see Kubozono, this volume). For this reason, the phonetic properties of affricate geminates have been much understudied. As far as I know, the only extensive study is that is offered by Oba, Brown, and Handke (2009), who found that the primary acoustic correlate of affricate geminates seems to lie in the difference in the closure duration, and not in frication duration. More studies on the properties of affricate geminates in Japanese are hoped for.

### 2.3.3 Voiced obstruent geminates

Finally, the effect of voicing on geminates is no less interesting. The native phonology of Japanese does not allow voiced obstruent geminates (Ito and Mester 1995, 1999; Kuroda 1965). The lack of voiced obstruent geminates has been argued to be due to their aerodynamic difficulty (Hayes and Steriade 2004; Ohala 1983; Westbury and

---

**10** This study also found that the duration ratio of [p]-[pp] is smaller than that of [t]-[tt] and [k]-[kk]. This lower ratio may be related to the fact that length is not contrastive for [p] in the native phonology in Japanese (see Ito and Mester 1995, 1999 and Nasu, this volume). One puzzle, however, is why voiced stops have high duration ratios despite the fact that they are not contrastive in native Japanese phonology (Ito and Mester 1995, 1999). See also Engstrand and Krull (1994) for the relationship between the functional load of length contrasts and their phonetic realization. A full consideration on this relationship should be explored in future studies.

**11** Whether there indeed is a difference in perceptibility between stops and fricatives should be tested in a perception study.

Keating 1986, and more references cited in Kawahara 2006a). For voiced stops, the intraoral air pressure goes up behind the oral stop closure; this rise in the intraoral air pressure makes it difficult to maintain the airflow required for vocal fold vibration. For voiced fricatives, the intraoral airpressure must rise to create frication, which again makes it difficult to maintain the transglottal air pressure drop. Perhaps for these reasons (synchronically or diachronically), the native phonology of Japanese does not allow voiced obstruent geminates.

However, gemination found in the context of loanword adaptation resulted in voiced obstruent geminates (e.g. Katayama 1998; Kubozono, Ito, and Mester 2008; Shirai 2002; see also Kawagoe, this volume, and Kubozono, this volume); e.g. [heddo] 'head' and [eggɯ] 'egg'. Nevertheless, presumably due to the aerodynamic difficulty, voiced geminate stops are generally "semi-devoiced" in Japanese. All three speakers recorded in Kawahara (2006a) show semi-devoicing. Figures 3 and 4 illustrate the difference between singletons and geminates: for singleton [g], closure voicing is fully maintained, while for geminate [gg], voicing during the stop closure ceases in the middle of the whole closure.[12] In Kawahara (2006a), on average, voicing is maintained only about 40% of the whole closure. Hirose and Ashby (2007) replicate this finding, showing that voiced Japanese geminates have only 47% of closure voicing.

As far as I know, there is no quantitative study on the phonetic implementation on voiced geminate fricatives in Japanese – this is a topic which is worth pursuing in a future study.[13]

One notable aspect of this semi-devoicing of geminates is that the following word-final high vowels after "semi-devoiced" geminates (e.g. [eggɯ̥] 'egg') do not devoice, even though the vowels are preceded by a – phonetically speaking – voiceless interval (Hirose and Ashby 2007). The lack of high vowel devoicing in this context shows that the (semi-devoiced) voiced geminates are still phonologically voiced, and that high vowel devoicing is conditioned by phonological, rather than, phonetic factors. See Fujimoto (this volume) for further discussion on this debate.

The semi-devoicing of voiced obstruent geminates is found in other languages (e.g. (Tashlhiyt) Berber: Ridouane 2010), but it is not universal, despite the fact that it presumably arises from a physical, aerodynamic difficulty (Ohala 1983). Cohn, Ham and Podesva (1999) show, for example, that Buginese, Madurese, and Toba Batak all maintain voicing throughout the geminate closure; Egyptian Arabic is another language which has fully voiced geminates (Kawahara 2006a), and Lebanese

12 These spectrograms are based on new recordings made for Kawahara (2013c).

13 Voiced fricatives in Japanese become affricates word-initially, although whether this alternation is in free-variation or an allophonic alternation is controversial (Maekawa 2010). Osamu Fujimura (p.c., April 2012) points out that this hardening process may also happen when voiced fricatives become geminates as well, as in [oddzɯ] 'odds'. Affrication process may then be a general hardening process, which occurs in phonetically strong positions (i.e. word-initially and in geminates).

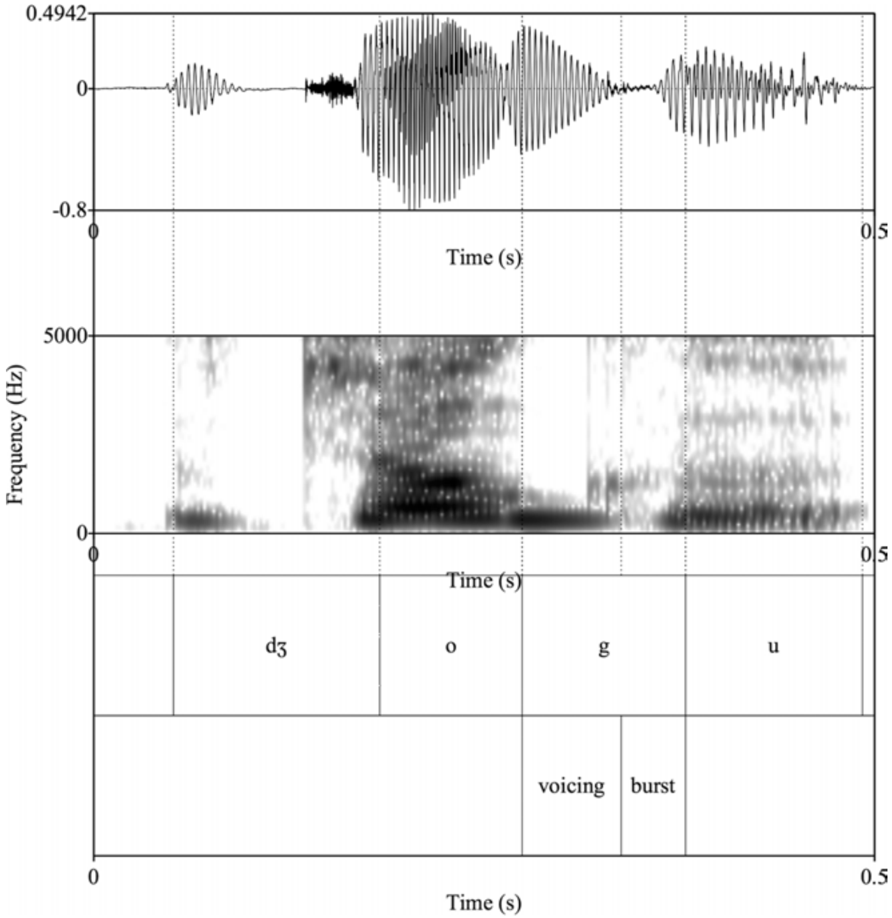


**Figure 3:** A singleton [g]

Arabic shows high percentages of voicing maintenance in medial, non-final positions (Ham 2001). Several Japanese dialects in Kyushu, including the Nagasaki dialect, also seem to show fully voiced geminate stops (Matsuura 2012). Cohn, Ham, and Podesva (1999) speculate that speakers resort to extra articulatory maneuvers like larynx lowering and cheek expansion to deal with the aerodynamic challenges (Ohala 1983). These articulatory gestures expand the size of oral cavity, thereby lowering the intraoral pressure (by Boyle's Law), providing the sufficient transglottal air pressure drop necessary to maintain vocal fold vibration (see Hayes and Steriade 2004, Ohala 1983, Ohala and Riordan 1979, and others).

The reason that (non-Kyushu) Japanese speakers do not deploy such articulatory strategies – at least not to the extent that geminates are fully voiced – may be that

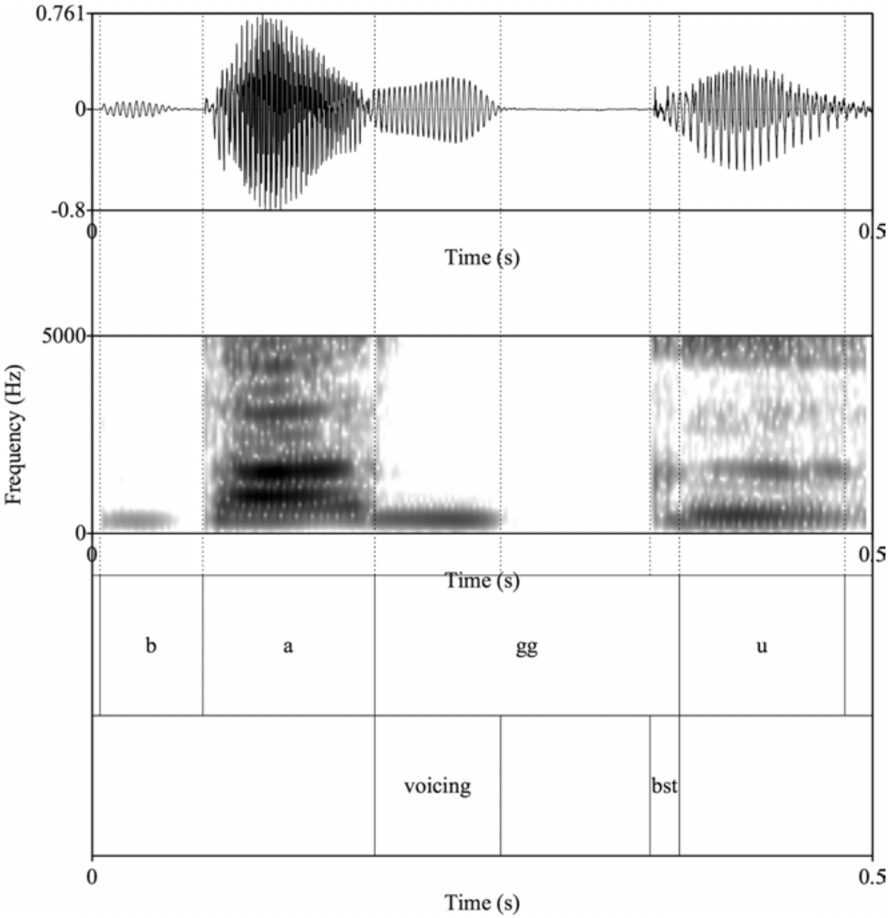


**Figure 4:** A geminate [gg]

voiced obstruent geminates are historically relatively new (see Pintér, this volume), and therefore the functional load of a voicing contrast in geminates is low, the contrast being restricted to loanwords (Ito and Mester 1995, 1999); i.e. there are not many minimal pairs. It would thus be interesting to observe whether speakers of future generations would start producing fully-voiced geminates, if the voicing contrast in geminate becomes more widespread in the Japanese lexicon. Moreover, a further cross-linguistic comparison is warranted to explore the relationship between how voiced stop geminates are implemented, and how the particular phonetic implementation patterns affect their phonological patterns (if they do at all) (see Kawahara 2006a for discussion).

## 2.4 Comparison with other languages

### 2.4.1 Constriction duration

I have already mentioned a few differences and similarities between Japanese geminates and geminates found in other languages, but now we turn our attention to a more detailed comparison of Japanese with other languages. As reviewed in section 2.1, Japanese geminates are acoustically characterized by long constriction duration, almost always twice as long as corresponding singletons. Similarly, constriction duration is usually the primary acoustic correlate of a singleton/geminate contrast in other languages; e.g. (Lebanese) Arabic (Ham 2001), Bengali (Lahiri and Hankamer 1988), Berber (Ridouane 2010), Bernese (Ham 2001), Buginese (Cohn, Ham, and Podesva 1999), Estonian (Engstrand and Krull 1994), Finnish (Engstrand and Krull 1994; Lehtonen 1970), Cypriot Greek (Tserdanelis and Arvaniti 2001), Guinaang Bontok (Aoyama and Reid 2006), Hindi (Ohala 2007; Shrotriya et al. 1995), Hungarian (Ham 2001), Italian (Esposito and Di Benedetto 1999; Payne 2005; Pickett, Blumstein, and Burton 1999), Madurese (Cohn, Ham, and Podesva 1999), Malayalam (Local and Simpson 1999), Pattani Malay (Abramson 1987b), Persian (Hansen 2004), Swedish (Engstrand and Krull 1994), Swiss German (Kraehenmann and Lahiri 2008), Toba Batak (Cohn, Ham, and Podesva 1999), and Turkish (Lahiri and Hankamer 1988) (see Kawahara and Braver 2014 and Ridouane 2010 for more languages and references).

One interesting cross-linguistic difference is the size of duration ratios between singletons and geminates. In Norwegian, for example, the ratio is much smaller than in Japanese (ranging from 1.22–1.38 in medial positions, cf. Table 1), and more substantial differences manifest themselves in the duration of preceding vowels (Fintoft 1961) (although one should note that Fintoft measured only non-stop consonants; see section 2.3.1).[14] In Buginese and Madurese, the geminate/singleton duration ratios are generally below 2 (Cohn, Ham, and Podesva 1999). Generalizing this observation, Ham (2001) entertains the possibility that geminate/singleton duration ratios are smaller for syllable-timed languages (e.g. Norwegian) than for mora-timed languages (e.g. Japanese). See also Maekawa (1984) for a comparison between Standard Tokyo dialect and Akita dialect – a dialect that has been described as syllable-timed – which points to the same generalization.

### 2.4.2 Other durational correlates

As discussed in section 2.2.1, vowels are longer before geminates in Japanese. This observation may come as a surprise given a cross-linguistic tendency that vowels in

**14** Accordingly, when perceiving a singleton/geminate contrast, Norwegian speakers substantially rely on preceding vowel duration, much more than speakers of other languages (Kingston et al. 2009).

closed syllables are often shorter than vowels in open syllables (Maddieson 1985). Indeed many languages have shorter vowels before geminates than before singletons; e.g. Bengali (Lahiri and Hankamer 1988), Berber (Ridouane 2010), Italian (Esposito and Di Benedetto 1999; Pickett, Blumstein, and Burton 1999), Hindi (Ohala 2007; Shrotriya et al. 1995), Malayalam (Local and Simpson 1999), and the three Polynesian languages studied by Cohn, Ham, and Podesva (1999).

However, there are other languages that arguably show lengthening of vowels before geminates: Turkish,[15] Finnish (Lehtonen 1970, pp. 110–111), Shinhala (Letterman 1994) (although only one of the two speakers showed clear evidence) and Persian (Hansen 2004) (although no direct statistical tests are reported). The existence of such languages shows that Japanese may not simply be a typological anomaly, but languages vary in whether geminates shorten or lengthen the preceding vowels. I will come back to this issue of this cross-linguistic difference in section 3.2 in relation to its perceptual relevance.

In some languages, there are no substantial differences in the preceding vowel duration with singletons and geminates; e.g. Egyptian Arabic (Norlin 1987), Lebanese Arabic (at least for short vowels) (Ham 2001), Estonian (Engstrand and Krull 1994), and Hungarian (Ham 2001). In Cypriot Greek, there is slight tendency toward shortening before geminates, but this tendency is not very consistent (Tserdanelis and Arvaniti 2001).

Finally, the lack of an effect of geminacy on VOT in Japanese is paralleled in many languages including Buginese, Madurese, Toba Batak (Cohn, Ham, and Podesva 1999), Bernese, Hungarian, Lebanese Arabic (Ham 2001), Bengali (Hankamer, Lahiri, and Koreman 1989), and Berber (Ridouane 2010). Cypriot Greek has consistently longer VOT for geminates (Tserdanelis and Arvaniti 2001), but Turkish shows shorter VOT for geminates (Lahiri and Hankamer 1988).

### 2.4.3 Other non-durational, acoustic correlates

In addition to the durational correlates, different languages seem to show different non-durational acoustic correlates to signal singleton-geminate contrasts. These non-durational correlates are summarized in (1)–(6).[16]

(1) Bengali (Hankamer, Lahiri and Koreman 1989)
   a. Root Mean Square (RMS) amplitude of the following syllable is higher after singletons.

15 In Lahiri and Hankamer (1988), the difference is small and not statistically significant; see also Jannedy (1995) for evidence that this lengthening applies to closed syllables in general, as in Japanese (see footnote 4).

16 See the original references for stimulus designs and measurement procedures.

(2) Berber (Ridouane 2010)
  a. Geminates have higher amplitude during release.
  b. Geminates show burst release more consistently than singletons.

(3) Hindi (Shrotriya et al. 1995)
  a. F0 rises toward geminates in the preceding vowel.
  b. Burst intensity is stronger for geminates (by about 10dB).

(4) Italian (Payne 2006, based on electropalatographic (EPG) data)
  a. Geminates involve a more palatalized constriction than singletons.
  b. Geminate stops involve a more complete occlusion.
  c. Geminates are associated with a laminal gesture; singletons are associated with an apical gesture.

(5) Malayalam (Local and Simpson 1999)
  a. Sonorant geminates show palatal resonance with higher F2.
  b. The surrounding vowels differ in F1 and F2.

(6) Pattani Malay
  a. The peak amplitude of initial vowels (with respect to the following vowel) is higher after word-initial geminates than singletons (Abramson 1987b, 1998).
  b. Fundamental frequency of word-initial vowels is higher after word-initial geminates (Abramson 1998).
  c. First vowels are longer (with respect to second vowels) after word-initial geminates (Abramson 1998).
  d. The slope of amplitude rise is steeper after word-initial geminates (Abramson 1998).

So far Idemaru and Guion (2008) is the most extensive study looking for spectral correlates of geminacy contrasts in Japanese, and it is yet to be investigated whether the correlates listed in (1)–(6) are found in Japanese (though the Pattani Malay case may be special because it involves cases of word-initial geminates). However, it seems likely at this point that the phonetic implementation patterns of singleton-geminate contrasts are language-specific, the only universal rule being that geminates are longer than singletons (Ham 2001; Ridouane 2010). A remaining task in the phonetic theory is how to model the universality and language-specificity of phonetic implementation patterns of length contrasts. We should also perhaps bear in mind that "geminates" in different languages may not be the same phonological entity – there remains a possibility that these "geminates" have different phonological representations. See also Davis (2011) for relevant discussion.

# 3 The perception of geminates in Japanese

We now turn to the perception of a singleton-geminate contrast, beginning with a discussion of cues used by Japanese listeners and continuing with a discussion of cross-linguistic cues for geminacy contrasts.

## 3.1 The primary cue: constriction duration

Many studies have shown that the longer the constriction, the more likely the target is perceived as a geminate. This effect has been shown to hold in many perception studies using Japanese listeners (Amano and Hirata 2010; Arai and Kawagoe 1998; Fujisaki, Nakamura, and Imoto 1975; Fujisaki and Sugito 1977; Fukui 1978; Hirata 1990; Kingston et al. 2009; Oba, Brown, and Handke 2009; Takeyasu 2012; Watanabe and Hirato 1985). As an example, Figure 5 reproduces the results of Kingston et al.

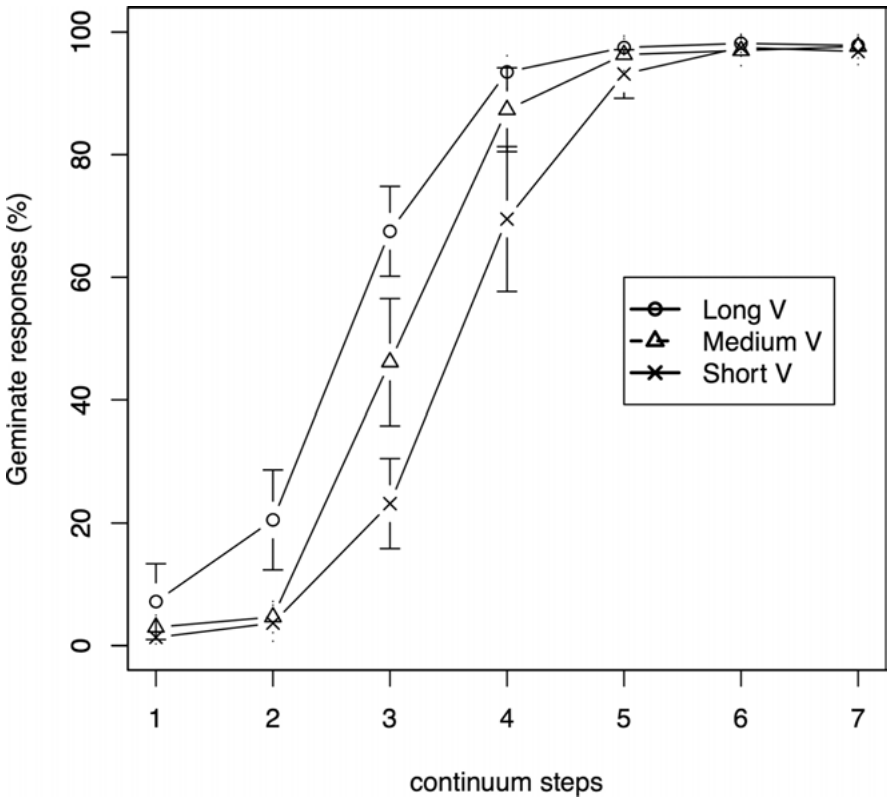


**Figure 5:** The effect of closure duration and the preceding vowel duration on the perception of geminates by Japanese listeners. Adapted from Kingston et al. (2009). Reprinted with permission from Elsevier

(2009) in which closure duration was varied from 60ms and 150ms in 15ms increments (see the next section for the three vocalic contexts). We observe that geminate responses increase as closure duration increases.

## 3.2 Contextual effects

More controversial than the effects of constriction duration are contextual effects. Fukui (1978) found that when the closure duration of an original singleton consonant was lengthened, it was almost always perceived as a geminate when the closure duration was doubled. On the other hand, shortening an original geminate did not result in a comparable shift in perception. The results show that closure duration is not the only cue for perceiving geminates. Similar types of effects (albeit to different degrees) were found in similar types of experiments on other languages (Bengali: Hankamer, Lahiri, and Koreman 1989, Pattani Malay: Abramson 1987a, 1992, Tamil: Lisker 1958, and Turkish: Hankamer, Lahiri, and Koreman 1989).

As reviewed in section 2.2.1, given that vowels are longer before geminates, we expect that Japanese speakers are more likely to perceive a consonant as a geminate after a longer vowel than after a shorter vowel. Several results indeed found a contextual effect in this direction (Arai and Kawagoe 1998; Kingston et al. 2009; Ofuka 2003; Ofuka, Mori, and Kiritani 2005; Takeyasu 2012). This contextual effect is illustrated in Figure 5 in which listeners judged more of the continuum as geminates after longer vowels.

On the other hand, several studies have found opposite results as well. For example, Watanabe and Hirato (1985) found that the perceptual boundaries between singletons and geminates shift toward longer duration after longer vowels, meaning that longer duration was required after longer vowels for consonants to be perceived as geminates (although only two listeners participated in this study). A similar boundary shift was found in Hirata (1990). Idemaru and Guion-Anderson (2010) kept the duration of the consonant at about 140ms and changed the duration of the preceding mora ($C_1 + V_1$ = onset plus preceding vowel), and found that the shorter the preceding mora duration, the more geminate responses were obtained. On the other hand, Takeyasu (2012) argues that it is the duration of $C_1/V_1$ ratio that matters, and that higher $C_1/V_1$ ratios lead to more geminate percepts. For more references of studies that obtained the results in this direction, see also Fujisaki and Sugito (1977)[17] and Idemaru, Holt, and Seltman (2012).

In summary, some studies found an "assimilative" pattern (more geminate responses after longer vowels) while others found a "contrastive" pattern (more geminate responses after shorter vowels). Where the difference between the two types

**17** Fujisaki and Sugito (1977) found a contextual effect for the /s/-/ss/ contrast, but the paper is not explicit about the other two geminate pairs (/t/-/tt/ and /m/-/mm/).

of results comes from is an interesting question. There is some evidence that the magnitudes of the duration ratios between the target and context may matter in this regard (Nakajima, ten Hoopen, and Hilkhuysen 1992). Takeyasu (2012) also entertains the hypothesis that in experiments that obtained an contrastive effect, listeners may have judged the preceding vowels to be phonologically long, in which case the listeners are biased against judging the following consonant as long to avoid a superheavy syllable (see Ito and Mester, this volume, and Kubozono 1999 for a phonological constraint against superheavy syllables in Japanese, and Kawagoe and Takemura 2013 for its perceptual impact). Further experimentation is necessary to settle this issue.

Unlike preceding vowels, vowels are shorter after geminates than after singletons (Campbell 1999; Han 1994; Idemaru and Guion 2008; Ofuka 2003) (see section 2.2.1). While Hirato and Watanabe (1987) found no effects of the duration of the following vowel on the perception of geminates, Ofuka, Mori and Kiritani (2005) did in fact find that listeners are more likely to judge stimuli as a geminate before a shorter vowel; Idemaru and Guion-Anderson (2010) found a similar contextual effect of following vowels, although they found the effect of preceding $C_1V_1$ mora to be more substantial. See also Nakajima, ten Hoopen, and Hilkhuysen (1992) for a relevant discussion.

Another issue is the (non-)locality of contextual effects. For example, Hirata (1990) tested the effect of sentence level speech rate on perception of length contrasts, and found that the duration of the whole sentential materials following the target word can impact the perception of geminates. The study found that those tokens which are unambiguously identified as either a singleton or a geminate can be perceived as a member of a different category if the following materials provide enough cues for speech rate.

When listeners normalize the perceived duration for speech rate, one remaining question is: to what extent do they rely on local cues like the immediately preceding/following vowels or (CV) moras or (C)VC(C)V subword (Hirata and Amano 2012), and to what extent do they rely on more global cues (like the entire word or utterance). On the one hand, in terms of psycholinguistic computational simplicity, local cues are presumably easier to track (Idemaru and Guion-Anderson 2010). Nevertheless, some studies (Amano and Hirata 2010; Hirata 1990; Pickett and Decker 1960) show the effect of global cues; for example, by comparing several relational measures, Amano and Hirata (2010) demonstrate that the relationship between consonant duration and entire word duration[18] provides a good perceptual cue to a length distinction in Japanese. Recall also that Hirata (1990) found contextual effects at sentential levels.

**18** They demonstrate that it is not a simple ratio between these two durations, but a regression function with an intercept that most accurately predicts the perceptual behavior of Japanese listeners. This function is equivalent to the ratio between closure duration ($c$) plus some constant ($k$) and word duration ($w$); i.e. $(c + k)/w$.

However, taking into account a whole word or sentence to determine a length property of a singleton/geminate contrast may impose a psycholinguistic burden. In order to identify what the word is, it is necessary to determine whether the consonant in question is a singleton or a geminate, but in order for listeners to determine whether the consonant is singleton or a geminate, they need to know what the word is – there may be a chicken-and-egg problem here.

I do not wish to imply that this challenge is insurmountable, rather that more phonetic and psycholinguistic research is necessary to address this issue. Hirata (2007) suggests that gating experiments (Grosjean 1980) may address the issue of the (non-)locality of the perception of length contrasts. In this way, the relationship between production and perception of geminates in Japanese (as well as in other languages) provides an interesting forum of research, which may bear on the general theory of speech perception (see Amano and Hirata 2010, Hirata and Amano 2012, Idemaru and Guion-Anderson 2010, Idemaru, Holt, and Seltman 2012, Otaki 2011, Pind 1986: and others for discussion).

Another remaining question is how non-durational cues – F0 values and movement, spectral envelope, burst intensity, etc. (see also Table 2) – interact with durational cues in the perception of Japanese geminates. For example, Ofuka (2003) observes that geminates are shorter in accented disyllabic words than in corresponding unaccented words, and also that in perception, a consonant with a particular duration is more likely to be perceived as a geminate when the word is accented (see also Hirata 1990 who obtained similar results). Likewise, Kubozono, Takeyasu, and Giriko (2013) show that English monosyllabic utterances with falling pitch contours – which are acoustically similar to Japanese pitch accents (Kawahara, this volume) – are more likely to be perceived as geminates by Japanese listeners. On the other hand, Idemaru (2011) did not find any substantial effects of amplitude or the steepness of F0 fall on the perception of geminacy for Japanese listeners. More extensive studies are warranted to investigate the intricacy of perception of geminates in Japanese.

## 3.3 Comparison with other languages

Like Japanese, the effect of constriction duration on the perception of duration has been found in many languages; e.g. Arabic (Obrecht 1965), Bengali (Hankamer, Lahiri, and Koreman 1989), English[19] (Pickett and Decker 1960), Finnish (Lehtonen 1970), Hindi (Shrotriya et al. 1995), Italian (Esposito and Di Benedetto 1999; Kingston et al. 2009), Norwegian (Kingston et al. 2009), Pattani Malay (Abramson 1987a, 1992), and Turkish (Hankamer, Lahiri, and Koreman 1989).

**19** English does not have a lexical geminate contrast; this experiment tested a pair like *topic* vs. *top pick* where one member of the pair contains multiple morphemes.

Across languages, the effect of a language particular phonetic implementation pattern – shortening or lengthening of the preceding vowel – is often reflected in the perception pattern as well. For example, unlike in Japanese, in both Norwegian (Fintoft 1961) and Italian (Esposito and Di Benedetto 1999), vowels are shorter before geminates. This shortening affects the perception of geminates – listeners of these languages are more likely to perceive a consonant as a geminate before a shorter vowel than a longer vowel (Esposito and Di Benedetto 1999; Kingston et al. 2009; van Dommelen 1999). In Icelandic, in which long vowels and geminates are in a complementary distribution, Pind (1986) shows that vowel duration with respect to the entire rhyme duration is a good predictor of geminate perception – given fixed rhyme durations, shorter vowel durations yielded more geminate responses.

One interesting puzzle that arises from this cross-linguistic comparison regarding shortening vs. lengthening in pre-geminate position is as follows: some researchers propose that C/V duration ratios provide mutually enhancing perceptual cues for duration when a shorter consonant is preceded by a longer vowel, as is the case for voicing contrasts in many languages (Kingston and Diehl 1994; Kohler 1979; Pickett, Blumstein, and Burton 1999; Port and Dalby 1982). A combination of a short vowel and a long consonant yields enhanced, high $C/V_1$ duration ratios, whereas a combination of a long vowel and a short consonant yields low ratios. Languages like Italian and Norwegian, in which preceding vowels are shorter before geminates, can be assumed to deploy this perceptual enhancement pattern. In this light, a question arises why Japanese lengthens a vowel before a geminate.

A tentative answer that I can offer is that $V_1C$ unit (or V-to-V interval) may constitute another kind of perceptual unit, a unit that has been hypothesized to play a role in the perception of Japanese and other languages (Hirata and Forbes 2007; Kato, Tsuzaki, and Sagisaka 2003; Kingston et al. 2009; Ofuka, Mori, and Kiritani 2005; Sato 1978; van Dommelen 1999).[20] If $V_1C$ is an important perceptual unit – whether it is universal or specific to Japanese – then a longer vowel before a geminate can be considered as perceptually enhancing the longer duration of geminates.

## 4 The articulatory characteristics of Japanese geminates

Compared to acoustic and perception studies of Japanese geminates, there are relatively fewer studies on the articulation of Japanese geminates, although there

20 An alternative idea is that although Japanese is a mora-timed language (where a mora usually constitutes a CV unit), geminates, whose coda part should constitute its own mora, are not by themselves as long as a CV unit; pre-geminate vowel lengthening may occur to compensate for this shortage of duration, as hypothesized and discussed by Warner and Arai (1999). See also Otake (this volume) for more on mora-timing in Japanese. One puzzle for this explanation is why, then, Japanese speakers shorten the following vowels after geminates.

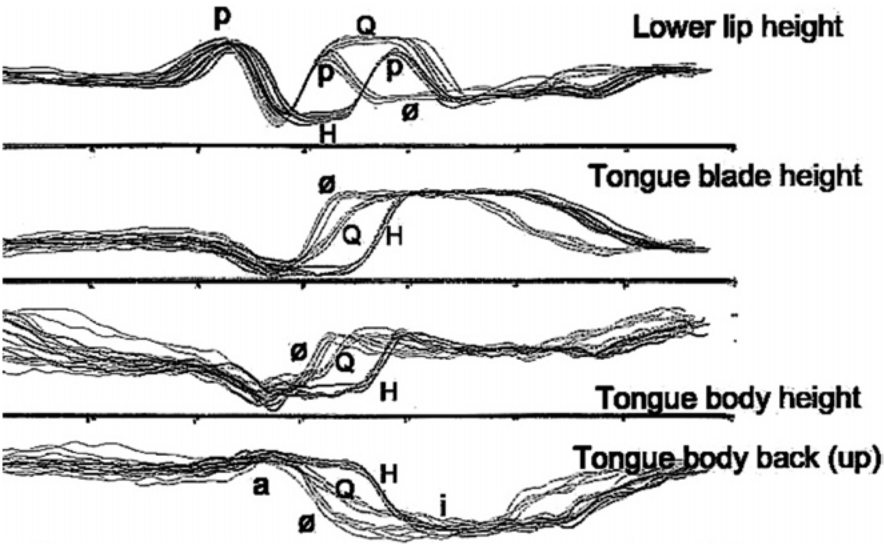


**Figure 6:** The articulatory movements of Japanese geminates, as compared to singletons. Based on Ishii (1999), cited and discussed in Fujimura & Williams (2008). Three conditions are [papa] (ɸ), [paapa] (H), and [pappa] (Q). Reprinted with permission from the author and the publisher

are some notable studies. Ishii (1999), for example, obtained articulatory data of Japanese geminates and long vowels using X-ray microbeam measurements, as shown in Figure 6. The three types of the stimuli were tested in this study, which were [papa] (ɸ), [paapa] (H), and [pappa] (Q).

Based on Figure 6, Fujimura and Williams (2008) make three observations. First, as we can observe in the top panel, a geminate [pp] in Japanese shows a prolonged lip closure compared to a singleton [p]. Second, while the lip movement toward its closure is comparable between singletons and geminates (the top panel) (though cf. Löfqvist 2007; Smith 1995), the lingual (tongue) movements are slower for geminates than for singletons (the second and the third panel). Finally, the V-to-V movement is slower and more gradual across geminates than across singletons (the bottom panel).

These results are corroborated by studies by Löfqvist (2006, 2007) using a magnetometer system. Longer constriction duration was confirmed for labial (Löfqvist 2006)[21] as well as alveolar and velar stops (Kochetov 2012; Löfqvist 2007). The speed of the tongue movement was found to be slower for alveolar and velar geminate stops than corresponding singletons (Löfqvist 2007). Slower V-to-V movement across geminate stops was also found by Löfqvist (2006).

**21** Löfqvist (2006) studies nasal geminates, and therefore this finding is technically for *hatsuon*, not for *sokuon*.

Takada (1985) investigated X-ray data of Japanese consonants, and found two differences between singletons and geminates: slower movement in terms of lingual contact and jaw contact, with maximal contact formed at a later phase in the constriction in geminates. Smith (1995), again based on X-ray microbeam data, shows that a singleton/geminate distinction affects the gestural timing of the following vowel in Japanese, whereas in Italian it does not – she attributes this difference to differences in gestural coordination of vowels and consonants in Japanese and Italian. The EPG data by Kochetov (2012) shows greater degree of linguopalatal contact for geminates than for singletons. Sawashima (1968), using a fiberscope, shows that glottal abduction is larger for geminate fricatives than singleton fricatives. Finally, Kokuritsu Kokugo Kenkyūjo (1990) offers detailed articulatory data of Japanese sounds in general, including those of geminates.

# 5 Remaining issues

Although I have raised a number of remaining questions already, I would like to close this chapter with a discussion about several more questions that require further experimentation.

## 5.1 Non-intervocalic geminates

For lexical contrasts, Japanese allows geminates only intervocalically. However, some word-initial geminates are found due to an elision process in casual speech; e.g. [ttakɯ] from /mattaku/ (a phrase that often accompanies a sigh) and [sseena] from /usseena/ 'shut up'. Cues to word-initial geminates have been studied in some other languages (Abramson 1992, 1999; Kraehenmann and Lahiri 2008; Kraehenmann 2011; Muller 2001; Ridouane 2010), but the Japanese case has not been extensively investigated. A specific question is whether such word-initial geminates involve longer constriction just like intervocalic geminates. Articulatory studies, using devices like EPG (Kraehenmann and Lahiri 2008; Payne 2006; Ridouane 2010), would address the question of whether geminates do indeed involve a longer constriction word-initially (see Kraehenmann and Lahiri 2008 and Ridouane 2010 who found a positive answer to this question in Swiss German and Berber).

Similarly, an orthographic marker for Japanese geminates – "small tsu" – can also appear word-finally, especially in mimetic words (see Nasu, this volume), although this word-final gemination diacritic does not convey a lexical contrast. The exact nature of its phonetic realization is yet to be explored – impressionistically, it is realized as a glottal stop, but as far as I know, it has not been fully explored in instrumental work.

## 5.2 Derived geminates vs. underlying geminates

Some phonetic studies in other languages have compared lexical geminates and geminates derived by some phonological processes, most often by assimilation.[22] They have generally shown that lexical geminates and geminates derived by phonological processes are phonetically identical, as in Bengali (Lahiri and Hankamer 1988), Berber (Ridouane 2010), Sardinian (Ladd and Scobbie 2003), and Turkish (Lahiri and Hankamer 1988). However, Ridouane (2010) found a difference between lexical geminates and geminates created via morpheme concatenation in terms of preceding vowel duration and burst amplitude. Similarly, Payne (2005) argues that in Italian lexical geminates tend to be longer than post-lexical geminates created by RADDOPPIAMENTO SINTATTICO (RS) (although there are some complicating factors; see Payne 2006 for further discussion).

As far as I know, no studies have compared underlying and derived geminates in Japanese. For example, the final consonant of a prefix /maC-/ 'truly' assimilates to the root-initial consonant, resulting in a geminate (e.g. [mak-ka] 'truly red', [mas-sakasama] 'truly reversed', and [mam-marɯ] 'truly round'). It would be interesting to investigate whether there is a difference between such derived geminates and underlying geminates. One reason why we may expect a difference is as follows. Monomoraic roots in Japanese can be lengthened to have a long vowel, when pronounced in isolation without a case particle (Mori 2002); however, duration ratios between these lengthened vowels and short vowels are smaller than the ratios between underlying long vowels and short vowels found in the previous research – i.e., that this lengthening pattern is only incompletely neutralizing (Mori 2002 compares her results with the data from Beckman 1982 and Hoequist 1982; Braver and Kawahara 2014 confirmed that there are differences in duration between lengthened vowels and underlying long vowels within one experiment). It would be particularly interesting if we find such an incomplete neutralization pattern (Port and O'Dell 1985 *et seq.*)[23] in the context of gemination.

## 5.3 The phonetics of emphatic geminates

Japanese deploys gemination to convey emphatic meanings (e.g. [kattai] 'very hard' from [katai] 'hard') (Aizawa 1985; Kawahara 2001, 2006b, 2013b). In terms of orthography, this gemination can be written with multiple signs of gemination ("small tsu")

---

**22** In some languages, geminates arise via simple morpheme concatenation without a further phonological change (known as "fake geminates"); e.g. /pat̪ + t̪e/ → [pat̪t̪e] 'spread out (INFINITIVE)' in Bengali (Lahiri and Hankamer 1988). In Japanese, fake geminates rarely if ever arise because root-final consonants always assimilate to the following consonant anyway; i.e. fake geminates would not be distinguishable from assimilated geminates.

**23** For an overview of incomplete neutralizations, see Braver (2013), Kawahara (2011) and Yu (2011).

(Aizawa 1985). It would be interesting to investigate to what extent such repetition of geminate diacritics is reflected in actual production (and for that matter, can be tracked in perception). This issue is partly addressed by Kawahara and Braver (2014). A production study shows that at least some speakers can make a six-way duration differences, given five degrees of emphatic consonants (and non-emphatic consonants). Other speakers showed a steady correlation between emphasis levels and duration. The articulatory and perceptual properties of these emphatic geminates should be investigated more in future research.

Furthermore, this emphatic gemination pattern can create otherwise unacceptable types of geminates, such as voiced obstruent geminates in native words and approximant geminates (Aizawa 1985; Kawahara 2001; Kawahara and Braver 2014). Together with the general phonetic properties of emphatic geminates, the phonetic realization of approximant geminates in Japanese, in particular, is understudied and yet to be investigated.

## 5.4 The laryngeal "tension" of geminates

Despite the studies mentioned in section 4, the exact articulatory nature of Japanese geminacy contrasts is yet to be fully explored. One particular issue concerns whether Japanese geminates involve laryngeal constriction or not. Impressionistically, Japanese geminates are sometimes conceived of as having an accompanying glottal constriction. Hattori (1984) suggests that the first half of geminates involves glottal tension (p. 139). Aizawa (1985) uses a term "choked consonant" to refer to (emphatic) geminates. Idemaru and Guion (2008) also found shallower spectral tilt (H1-A1) in the vowels following geminates, indicating some creakiness, which implies some glottal constriction (although two other measures of creakiness did not show differences in their study). Fujimura and Williams (2008) argue that laryngealization is a distinctive characteristic of Japanese geminates, which may even contribute to the perception of geminates.

On the other hand, a study by Fujimoto, Maekawa, and Funatsu (2010) using a high-speed digital video recording system, did not find evidence for laryngeal or glottal tension in Japanese geminates. They also found that glottal opening is slightly larger during (voiceless) geminates than during singletons. Therefore, whether Japanese geminates involve glottal tension, and if so how that glottalization is coordinated/synchronized with supralaryngeal (oral) gestures, is still to be explored.

## 5.5 Dialectal differences

There are few cross-dialectal studies on Japanese geminates, especially those written in English, which would be available to those scholars who do not read the Japanese literature. Due to the limitation of my expertise, I cannot discuss this issue extensively, but it would be particularly interesting to compare the properties of geminates in mora-timed dialects with syllable-timed dialects, such as the Aomori dialect (Takada 1985), the Akita dialect (Maekawa 1984), and the Kagoshima dialect (Kubozono and Matsui 2003).

### 5.6 Manner differences and the perception of geminates

Finally, as discussed in section 2.3, manner effects on the production of geminates in Japanese have been understudied. Relatedly, many perception experiments on Japanese geminates are based on voiceless stops (Amano and Hirata 2010; Arai and Kawagoe 1998; Hirata 1990; Hirato and Watanabe 1987; Fukui 1978; Idemaru and Guion-Anderson 2010; Kingston et al. 2009; Ofuka 2003; Takeyasu 2012; Watanabe and Hirato 1985). Fujisaki, Nakamura, and Imoto (1975) studied all manners, but nevertheless only report the results for fricatives (though see also Fujisaki and Sugito 1977 where they report the data for all manners). There are a few recent studies (Matsui 2012; Takeyasu 2009; Tews 2008), which investigated factors affecting the perception of geminates in fricatives. Oba, Brown, and Handke (2009) showed that the primary cue for affricate geminates lies in the closure phase, not in the frication phase. The production and the perception of different manners of geminates, including nasal geminates, warrants further investigation.

## Acknowledgements

Thanks to Megan Moran, Mel Pangilinan, and Akiko Takemura for collecting many of the references cited herein. For discussion on this paper and/or comments on earlier versions of this draft, I thank Osamu Fujimura, Manami Hirayama, Michinao Matsui, Toshio Matsuura, Hope McManus, Kazu Kurisu, Akiko Takemura, and two reviewers (Itsue Kawagoe and Yukari Hirata). Hope McManus reformatted and proofread this paper. Some portions of this paper have appeared in Kawahara (2013a), which is written in Japanese and additionally discusses aspects of long vowels. A part of this research is supported by the Research Council Grant from Rutgers University. Remaining errors are mine.

## References

Abramson, Arthur S. 1987a. The perception of word-initial consonant length: Pattani Malay. *Haskins Laboratories Status Report on Speech Research* SR-91. 149–156.

Abramson, Arthur S. 1987b. Word-initial consonant length in Pattani Malay. *Haskins Laboratories Status Report on Speech Research* SR-91. 143–147.

Abramson, Arthur S. 1992. Amplitude as a cue to word-initial consonant length: Pattani Malay. *Haskins Laboratories Status Report on Speech Research* SR-109/110. 251–254.

Abramson, Arthur S. 1998. The complex acoustic output of a single articulatory gesture: Pattani Malay word-initial consonant length. In Udom Warotamasikkhadit and Thanyarat Panakul (eds.), *Papers from the Fourth Annual Meeting of the Southeast Asian Linguistic Society 1994*, Tempe, Arizona: Arizona State University, 1–20.

Abramson, Arthur S. 1999. Fundamental frequency as a cue to word-initial consonant length: Pattani Malay. In John J. Ohala, Yoko Hasegawa, Manjari Ohala, Daniel Granville and Ashlee C. Bailey (eds.), *Proceedings of the 14th International Congress of Phonetic Sciences* (ICPhS 14), San Francisco: University of California, 591–594.

Abramson, Arthur S. and Leigh Lisker. 1985. Relative power of cues: F0 shift versus voice timing. In Victoria Fromkin (ed.), *Phonetic linguistics: Essays in honor of Peter Ladefoged*, 25–33. Orlando: Academic Press.

Aizawa, Yoshiko. 1985. Intensification by so-called "choked sounds"–long consonants–in Japanese. *The Study of Sounds* (The Phonetic Society of Japan) 21. 313–324.

Amano, Shigeaki and Yukari Hirata. 2010. Perception and production boundaries between single and geminate stops in Japanese. *Journal of the Acoustical Society of America* 128(4). 2049–2058.

Aoyama, Katsura and Lawrence Reid. 2006. Cross-linguistic tendencies and durational contrasts in geminate consonants: An examination of Bontok geminates. *Journal of International Phonetic Association* 36. 145–157.

Arai, Masako and Itsue Kawagoe. 1998. Eigo no onsetsu-gata to sokuon chikaku – Nansensugo ni yoru chikaku tesuto no hōkoku [Consonant gemination and syllable types in English: A perceptual test of nonsense words]. *Onsei Kenkyū* [Journal of the Phonetic Society of Japan] 2(3). 87–92.

Beckman, Mary. 1982. Segmental duration and the 'mora' in Japanese. *Phonetica* 39. 113–135.

Behne, Darn, Takayuki Arai, Peter Czigler and Kirk Sullivan. 1999. Vowel duration and spectra as perceptual cues to vowel quantity: A comparison of Japanese and Swedish. In John J. Ohala, Yoko Hasegawa, Manjari Ohala, Daniel Granville and Ashlee C. Bailey (eds.), *Proceedings of the 14th International Congress of Phonetic Sciences* (ICPhS 14), San Francisco: University of California, 857–860.

Blevins, Juliette. 2004. Klamath sibilant degemination: Implications of a recent sound change. *International Journal of American Linguistics* 70. 279–289.

Braver, Aaron. 2013. *Degrees of incompleteness in neutralization: Paradigm uniformity with a phonetics with weighted constraints*. New Brunswick, New Jersery: Rutgers University dissertation.

Braver, Aaron and Shigeto Kawahara. 2014. Incomplete vowel lengthening in Japanese: A first study. In Santana-Labarge Robert E. (ed.), *Proceedings of WCCFL 31.* Somerville: Cascadilla Press, 86–95.

Campbell, Nick. 1999. A study of Japanese speech timing from the syllable perspective. *Journal of the Phonetic Society of Japan* 3(2). 29–39.

Cohn, Abigail, William Ham and Robert Podesva. 1999. The phonetic realization of singleton-geminate contrasts in three languages of Indonesia. In John J. Ohala, Yoko Hasegawa, Manjari Ohala, Daniel Granville and Ashlee C. Bailey (eds.), *Proceedings of the 14th International Congress of Phonetic Sciences* (ICPhS 14), San Francisco: University of California, 587–590.

Davis, Stuart. 2011. Geminates. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume and Keren Rice (eds.), *The Blackwell companion to phonology 2*, 873–897. Malden, MA and Oxford: Wiley-Blackwell.

Diehl, Randy and Margaret Walsh. 1989. An auditory basis for the stimulus-length effect in the perception of stops and glides. *Journal of the Acoustical Society of America* 85. 2154–2164.

Diehl, Randy, Margaret Walsh and Keith Kluender. 1991. On the interpretability of speech/nonspeech comparisons: A reply to Fowler. *Journal of the Acoustical Society of America* 89(6). 2905–2909.

Engstrand, Olle and Diana Krull. 1994. Durational correlates of quantity in Swedish, Finnish and Estonian: Cross-language evidence for a theory of adaptive dispersion. *Phonetica* 51. 80–91.

Esposito, Anna and Maria Gabriella Di Benedetto. 1999. Acoustical and perceptual study of gemination in Italian stops. *Journal of the Acoustical Society of America* 106. 2051–2062.

Faluschi, Simone and Maria Gabriella Di Benedetto. 2001. Acoustic analysis of singleton and geminate affricates in Italian. *WEB-SLS: The European Journal of Language and Speech* (online).

Fintoft, Knut. 1961. The duration of some Norwegian speech sounds. *Phonetica* 7. 19–39.

Flemming, Edward. 2004. Contrast and perceptual distinctiveness. In Bruce Hayes, Rorbert Kirchner and Donca Steriade (eds.), *Phonetically based phonology*, 232–276. Cambridge: Cambridge University Press.

Fowler, Carol. 1990. Sound-producing sources as the objects of perception: Rate normalization and nonspeech perception. *Journal of the Acoustical Society of America* 88. 1236–1249.

Fowler, Carol. 1991. Auditory perception is not special: We see the world, we feel the world, we hear the world. *Journal of the Acoustical Society of America* 89. 2910–2915.

Fowler, Carol. 1992. Vowel duration and closure duration in voiced and unvoiced stops: There are no contrast effects here. *Journal of Phonetics* 20. 143–165.

Fujimoto, Masako. this volume. Chapter 4: Vowel devoicing. In Haruo Kubozono (ed.), *The handbook of Japanese phonetics and phonology.* Berlin: De Gruyter Mouton.

Fujimoto, Masako, Kikuo Maekawa and Seiya Funatsu. 2010. Laryngeal characteristics during the production of geminate consonants. In Takao Kobayashi, Keikichi Hirose and Satoshi Nakamura (eds.), *Proceedings of Interspeech 2010*, Chiba, 925–928.

Fujimura, Osamu and C. J. Williams. 2008. Prosody and syllables. *Phonological Studies* 11. 65–74.

Fujisaki, Hiroya, Kimie Nakamura and Toshiaki Imoto. 1975. Auditory perception of duration of speech and non-speech stimuli. In Gunnar Fant and Marcel A. A. Tatham (eds.), *Auditory analysis and perception of speech*, 197–219. London: Academic.

Fujisaki, Hiroya and Miyoko Sugito. 1977. Onsei no butsuriteki seishitsu [Physical natures of sounds]. In Susumu Ōno and Takesi Sibata (eds.), *Iwanami kōza nihongo 5: On'in* [Iwanami Japanese series 5: Phonology], 63–103. Tokyo: Iwanami.

Fukui, Seiji. 1978. Nihongo heisaon no enchō/tanshuku ni yoru sokuon/hisokuon to shite no chōshu [Perception for the Japanese stop consonants with reduced and extended durations]. *Onsei Gakkai Kaihō* [The Bulletin, The Phonetic Society of Japan] 159. 9–12.

Giovanardi, Maurizio and Maria Gabriella Di Benedetto. 1998. Acoustic analysis of singleton and geminate fricatives in Italian. *WEB-SLS: The European Journal of Language and Speech* (online).

Grosjean, Francois. 1980. Spoken word recognition processes and the gating paradigm. *Perception* & *Psychophysics* 28(4). 267–283.

Ham, William. 2001. *Phonetic and phonological aspects of geminate timing*. New York: Routledge.

Han, Mieko. 1962. The feature of duration in Japanese. *Onsei no Kenkyū* [The Study of Sounds, The Phonetic Society of Japan] 10. 65–80.

Han, Mieko. 1992. The timing control of geminate and single stop consonants in Japanese: A challenge for nonnative speakers. *Phonetica* 49. 102–127.

Han, Mieko. 1994. Acoustic manifestations of mora timing in Japanese. *Journal of the Acoustical Society of America* 96. 73–82.

Hankamer, George, Aditi Lahiri and Jacques Koreman. 1989. Perception of consonant length: Voiceless stops in Turkish and Bengali. *Journal of Phonetics* 17. 283–298.

Hansen, Benjamin B. 2004. Production of Persian geminate stops: Effects of varying speaking rate. In Augustine Agwuele, Willis Warren and Sang-Hoon Park (eds.), *Proceedings of the 2003 Texas Linguistics Society Conference*, 86–95. Somerville: Cascadilla Press.

Hattori, Shiro. 1984. *Onseigaku* [Phonetics]. Tokyo: Iwanami.

Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20. 253–306.

Hayes, Bruce and Donca Steriade. 2004. Introduction: The phonetic bases of phonological markedness. In Bruce Hayes, Robert Kirchner and Donca Steriade (eds.), *Phonetically based phonology*, 1–33. Cambridge: Cambridge University Press.

Hirata, Yukari. 1990. Tango reberu bun reberu ni okeru nihonjin no sokuon no kikitori [Perception of geminated stops in Japanese word and sentence levels]. *Onsei Gakkai Kaihō* [The Bulletin, The Phonetic Society of Japan] 194. 23–28.

Hirata, Yukari. 2004. Effects of speaking rate on the vowel length distinction in Japanese. *Journal of Phonetics* 32(4). 565–589.

Hirata, Yukari. 2007. Durational variability and invariance in Japanese stop quantity distinction: Roles of adjacent vowels. *Journal of the Phonetic Society of Japan* 11(1). 9–22.

Hirata, Yukari. this volume. Chapter 18: L2 phonetics and phonology. In Haruo Kubozono (ed.), *The handbook of Japanese phonetics and phonology.* Berlin: De Gruyter Mouton.

Hirata, Yukari and Connor Forbes. 2007. Vowel-onset interval as a timing unit for singleton/geminate stop distinction in Japanese. In Trouvain Jürgen and William J. Barry (eds.), *Proceedings of ICPhS XVI*, Saarbrücken, Germany, 349–354.

Hirata, Yukari and Stephen G. Lambacher. 2004. Role of word-external contexts in native speakers' identification of vowel length in Japanese. *Phonetica* 61. 177–200.

Hirata, Yukari and Kimiko Tsukada. 2009. Effects of speaking rate and vowel length on formant frequency displacement in Japanese. *Phonetica* 66(3). 129–149.

Hirata, Yukari and Jacob Whiton. 2005. Effects of speaking rate on the singleton/geminate distinction in Japanese. *Journal of the Acoustical Society of America* 118. 1647–1660.

Hirata, Yukari and Shigeaki Amano. 2012. Production of singleton and geminate stops in Japanese three- and four-mora Japanese words. *Journal of the Acoustical Society of America* 132(3). 1614–1625.

Hirato, Nobuo and Shinichiro Watanabe. 1987. Sokuon no chikaku to kōzoku boin no jizoku jikan tono kankei [The relationship between the perception of moraic obstruents and the duration of the following vowels]. *Onsei Gengo* II. 99–106.

Hirose, Aki and Michael Ashby. 2007. An acoustic study of devoicing of the geminate obstruents in Japanese. In Trouvain Jürgen and William J. Barry (eds.), *Proceedings of ICPhS XVI*, Saarbrücken, Germany, 909–912.

Hoequist, Carles E. 1982. Durational correlates of linguistic rhythm categories. *Phonetica* 40. 19–31.

Homma, Yayoi. 1981. Durational relationship between Japanese stops and vowels. *Journal of Phonetics* 9. 273–281.

Idemaru, Kaori. 2011. Role of amplitude and pitch in the perception of Japanese stop length contrasts. *Cross-cultural Studies* 24. 191–204.

Idemaru, Kaori and Susan Guion. 2008. Acoustic covariants of length contrast in Japanese stops. *Journal of International Phonetic Association* 38(2). 167–186.

Idemaru, Kaori and Susan Guion-Anderson. 2010. Relational timing in the production and perception of Japanese singleton and geminate stops. *Phonetica* 67. 25–46.

Idemaru, Kaori, Lori Holt and Howard Seltman. 2012. Individual differences in cue weights are stable across time: The case of Japanese stops lengths. *Journal of the Acoustical Society of America* 132. 3950–3964.

Ishii, Toyota. 1999. Sokuon hatsuwaji no kōon kikan no undō ni kansuru kenkyū [A study of the movement of articulatory organs in Japanese geminate production]. *Journal of Otolarngology of Japan* 102. 622–634.

Ito, Junko and Armin Mester. 1995. Japanese phonology. In John Goldsmith (ed.), *The handbook of phonological theory*, 817–838. Oxford: Blackwell.

Ito, Junko and Armin Mester. 1999. The phonological lexicon. In Natsuko Tsujimura (ed.), *The handbook of Japanese linguistics*, 62–100. Oxford: Blackwell.

Ito, Junko and Armin Mester. this volume. Chapter 9: Word formation and phonological processes. In Haruo Kubozono (ed.), *The handbook of Japanese phonetics and phonology*. Berlin: De Gruyter Mouton.

Jannedy, Stephanie. 1995. Gestural phasing as an explanation for vowel devoicing in Turkish. *Ohio State University Working Papers in Linguistics* 45. 56–84.

Katayama, Motoko. 1998. *Optimality Theory and Japanese loanword phonology*. Santa Cruz, California: University of California dissertation.

Kato, Hiroaki, Minoru Tsuzaki and Yoshinori Sagisaka. 2003. Functional differences between vowel onsets and offsets in temproal perception of speech: Local-change detection and speaking-rate discrimination. *Journal of the Acoustical Society of America* 113(6). 3379–3389.

Kawagoe, Itsue. this volume. Chapter 2: The phonology of sokuon, or geminate obstruents. In Haruo Kubozono (ed.), *The handbook of Japanese phonetics and phonology*. Berlin: De Gruyter Mouton.

Kawagoe, Itsue and Akiko Takemura. 2013. Geminate judgment of English-like words by Japanese native speakers: Differences in the borrowed forms of "stuff" and "tough". *Journal of East Asian Linguistics* 22(4). 307–337.

Kawahara, Shigeto. 2001. Similarity among variants: Output-variant correspondence. Tokyo: International Christian University BA thesis.

Kawahara, Shigeto. 2006a. A faithfulness ranking projected from a perceptibility scale: The case of [+voice] in Japanese. *Language* 82(3). 536–574.

Kawahara, Shigeto. 2006b. Mimetic gemination in Japanese: A challenge for Evolutionary Phonology. *Theoretical Linguistics* 32(1). 411–424.

Kawahara, Shigeto. 2011. Experimental approaches in theoretical phonology. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume and Keren Rice (eds.), *The Blackwell companion to phonology 4*, 2283–2303. Malden, MA and Oxford: Wiley-Blackwell.

Kawahara, Shigeto. 2012. Amplitude changes facilitate categorization and discrimination of length contrasts. *IEICE Technical Report. The Institute of Electronics, Information, and Communication Engineers* 112. 67–72.

Kawahara, Shigeto. 2013a. Nihongo tokushu haku no onkyō to chikaku [The acoustics and perception of Japanese special moras]. *Nihon Onkyō Gakkaishi* [Journal of the Acoustical Society of Japan] 69(4). 1–6.

Kawahara, Shigeto. 2013b. Emphatic gemination in Japanese mimetic words: A wug-test with auditory stimuli. *Language Sciences* 40. 300–307.

Kawahara, Shigeto. 2013c. Testing Japanese loanword devoicing: Addressing task effects. *Linguistics* 51. 1271–1299.

Kawhara, Shigeto. this volume. Chapter 11: The phonology of Japanese accent. In Haruo Kubozono (ed.), *The handbook of Japanese phonetics and phonology*. Berlin: De Gruyter Mouton.

Kawahara, Shigeto and Aaron Braver. 2013. The phonetics of emphatic vowel lengthening in Japanese. *Open Journal of Modern Linguistics* 3(2). 141–148.

Kawahara, Shigeto and Aaron Braver. 2014. The durational properties of emphatically-lengthened consonants in Japanese. *Journal of International Phonetic Association* 44(3). 237–260.

Kawahara, Shigeto, Melanie Pangilinan and Kelly Garvey. 2011. Spectral continuity and the perception of duration. Ms. Rutgers University. Talk presented at International Conference on Geminates 2011, Kobe University, 8–9 Jan.

Kingston, John and Randy Diehl. 1994. Phonetic knowledge. *Language* 70. 419–454.

Kingston, John, Shigeto Kawahara, Della Chambless, Daniel Mash and Eve Brenner-Alsop. 2009. Contextual effects on the perception of duration. *Journal of Phonetics* 37(3). 297–320.

Kinoshita, Keisuke, Dawn Behne and Takayuki. Arai. 2002. Duration and F0 as perceptual cues to Japanese vowel quantity. *Proceedings of ICSLP*, 757–760.

Kluender, Keith, Randy Diehl and Beverly Wright. 1988. Vowel-length differences before voiced and voiceless consonants: An auditory explanation. *Journal of Phonetics* 16. 153–169.

Kochetov, Alexei. 2012. Linguopalatal contact differences between Japanese geminate and singleton stops. *Canadian Acoustics* 40(3). 28–29.

Kohler, Klaus. 1979. Dimensions in the perception of fortis and lenis plosives. *Phonetica* 36. 332–343.

Kokuritsu Kokugo Kenkyūjo. 1990. *Nihongo no boin, shiin, onsetsu: Chōon undō no jikken onseigakuteki kenkyū* [Vowels, consonants and syllables in Japanese]. Tokyo: Shūei Shuppan.

Kraehenmann, Astrid and Aditi Lahiri. 2008. Duration differences in the articulation and acoustics of Swiss German word-initial geminate and singleton stops. *Journal of the Acoustical Society of America* 123(6). 4446–4455.

Kraehenmann, Astrid (2011) Initial geminates. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume and Keren Rice (eds.), *The Blackwell companion to phonology 2*, 1124–1146. Malden, MA and Oxford: Wiley-Blackwell.

Kubozono, Haruo. 1999. Mora and syllable. In Natsuko Tsujimura (ed.), *The handbook of Japanese linguistics*, 31–61. Oxford: Blackwell.

Kubozono, Haruo and Michinao Matsui. 2003. Phonetic vs. phonological control of speech: Closed syllable vowel shortening in Japanese dialects. In M. J. Solé, D. Recasens and J. Romero (eds.), *Proceedings of ICPhS XV*, 2429–2432.

Kubozono, Haruo, Junko Ito and Armin Mester. 2008. Consonant gemination in Japanese loanword phonology. In The Linguistic Society of Korea (ed.), *Current issues in unity and diversity of languages. Collection of papers selected from the 18th International Congress of Linguists*, 953–973. Republic of Korea: Dongam Publishing.

Kubozono, Haruo (ed.). 2007. *Sokuon*, or moraic obstruent. *Journal of the Phonetic Society of Japan* 11(1).

Kubozono, Haruo. this volume. Chapter 8: Loanword phonology. In Haruo Kubozono (ed.), *The handbook of Japanese phonetics and phonology*. Berlin: De Gruyter Mouton.

Kubozono, Haruo, Hajime Takeyasu and Mikio Giriko. 2013. On the positional assymmetry of consonant gemination in Japanese loanwords. *Journal of East Asian Linguistics* 22(4). 339–371.

Kuroda, S.-Y. 1965. *Generative grammatical studies in the Japanese language*. Cambridge, MA: MIT dissertation.

Ladd, D. Robert and James Scobbie. 2003. External sandhi as gestural overlap? Counter-evidence from Sardinian. In John Local, Richard Ogden and Rosalind Temple (eds.), *Papers in laboratory phonology VI: Phonetic interpretation*, 164–182. Cambridge: Cambridge University Press.

Lahiri, Aditi and George Hankamer. 1988. The timing of geminate consonants. *Journal of Phonetics* 16. 327–338.

Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge: MIT Press.

Lehtonen, Jaakko. 1970. *Aspects of quantity in standard Finnish*. Jyvaeskylaessae: Osakeyhtioen kirjapainossa.

Letterman, Rebecca S. 1994. A phonetic study of Sinhala syllable rhymes. *Working Papers of the Cornell Phonetics Laboratory* 9. 155–181.

Liljencrants, Johan and Björn Lindblom. 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48. 839–862.

Lindblom, Björn. 1986. Phonetic universals in vowel systems. In John Ohala and Jeri Jaeger (eds.), *Experimental phonology*, 13–44. Orlando: Academic Press.

Lisker, Leigh. 1958. The Tamil occlusives: Short vs. long or voiced vs. voiceless? *Indian Linguistics, Turner Jubilee* I. 294–301.

Lisker, Leigh. 1986. "Voicing" in English: A catalog of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech* 29. 3–11.

Local, John and Adrian Simpson. 1999. Phonetic implementation of geminates in Malayalam nouns. The phonetic realization of singleton-geminate contrasts in three languages of Indonesia. In John J. Ohala, Yoko Hasegawa, Manjari Ohala, Daniel Granville and Ashlee C. Bailey (eds.), *Proceedings of the 14th International Congress of Phonetic Sciences* (ICPhS 14), San Francisco: University of California, 595–598.

Löfqvist, Anders. 2006. Interarticulator programming: Effects of closure duration on lip and tongue coordination in Japanese. *Journal of the Acoustical Society of America* 120(5). 2872–2883.

Löfqvist, Anders. 2007. Tongue movement kinematics in long and short Japanese consonants. *Journal of the Acoustical Society of America* 122(1). 512–518.

Maddieson, Ian. 1985. Phonetic cues to syllabification. In Victoria Fromkin (ed.), *Phonetic linguistics*, 203–221. London: Academic Press.

Maekawa, Kikuo. 1984. Akita hōgen sokuon no jizoku jikan: "Sunzumari" no jittai to seiin [Duration of geminates in the Akita dialect]. *Hōgen Kenkyū Nenpō* 27. 231–247.

Maekawa, Kikuo. 2010. Coarticulatory reintrepretation of allophonic variation: Corpus-based analysis of /z/ in spontaneous Japanese. *Journal of Phonetics* 38(3). 360–374.

Martin, Andrew and Sharon Peperkamp. 2011. Speech perception and phonology. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume and Keren Rice (eds.), *The Blackwell companion to phonology 4*. 2334–2356. Malden, MA and Oxford: Wiley-Blackwell.

Matsuura, Toshio. 2012. Yūsei sogai jūshiin no onsei jitsugen ni okeru chiikisa ni kansuru yobiteki bunseki [A preliminary analysis on regional variation in phonetic realization of voiced obstruent geminates]. Talk presented at the Annual Convention of the Phonetic Society of Japan, 29 September.

Mattei, Marco and Maria Gabriella Di Benedetto. 2000. Acoustic analysis of singleton and geminate nasals in Italian. *WEB-SLS: The European Journal of Language and Speech* (online).

Matsui, Michinao. 2012. Shakuyōgo ni okeru sokuon seiki no yokusei yōin [An inhibitory factor of occurrence of voiceless geminate consonants in Japanese loanwords]. *Theoretical and Applied Linguistics at Kobe Shoin* 15. 49–102.

Miller, Joanne L. and Alvin M. Liberman. 1979. Some effects of later-occurring information on the perception of stop consonant and semivowel. *Perception & Psychophysics* 25. 457–465.

Mori, Yoko. 2002. Lengthening of Japanese monomoraic nouns. *Journal of Phonetics* 30(4). 689–708.

Motohashi-Saigo, Miki and Debra M. Hardison. 2009. Acquisition of L2 Japanese geminates: Training with waveform displays. *Language Learning & Training* 13(2). 29–47.

Muller, Jennier S. 2001. *The phonology and phonetics of word-initial geminates*. Columbus, Ohio: Ohio State University dissertation.

Nakajima, Yoshitaka, Gert ten Hoopen, and Gaston Hilkhuysen. 1992. Time-shrinking: A discontinuity in the perception of auditory temporal patterns. *Perception & Psychophysics* 51(5). 504–507.

Nasu, Akio. this volume. Chapter 6: The phonological lexicon and mimetic phonology. In Haruo Kubozono (ed.), *The handbook of Japanese phonetics and phonology*. Berlin: De Gruyter Mouton.

Norlin, Kjell. 1987. A phonetic study of emphasis and vowels in Egyptian Arabic. *Lund University Department of Linguistics Working Papers 30*.

Oba, Reiko, Angelika Braun and Jürgen Handke. 2009. The perception of Japanese geminates by native and non-native listeners. *ISPhS: International Society of Phonetic Sciences*, 9–29.

Obrecht, Dean. 1965. Three experiments in the perception of geminate consonants in Arabic. *Language and Speech* 8. 31–41.

Ofuka, Etsuko. 2003. Sokuon /tt/ no chikaku: Akusento-gata to sokuon/hisokuongo no onkyōteki tokuchō ni yoru chigai [Perception of a Japanese geminate stop /tt/: The effect of pitch type and acoustic characteristics of preceding/following vowels]. *Onsei Kenkyū* [Journal of the Phonetic Society of Japan] 7(1). 70–76.

Ofuka, Etsuko, Yoko Mori and Shigeru Kiritani. 2005. Sokuon no chikaku ni taisuru senkō kōzoku boinchō no eikyō [The effects of the duration of preceding and following vowels on the perception of geminates]. *Onsei Kenkyū* [Journal of the Phonetic Society of Japan] 9(2). 59–65.

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In Peter MacNeilage (ed.), *The production of speech*, 189–216. New York: Springer.

Ohala, John J. and Carol. J. Riordan. 1979. Passive vocal tract enlargement during voiced stops. In Jared. J. Wolf and Dennis. H. Klatt (eds), *Speech communication papers*, 89–92. New York: The Acoustical Society of America.

Ohala, Manjari. 2007. Experimental methods in the study of Hindi geminate consonants. In Maria Josep Solé, Patrice Beddor and Manjari Ohala (eds.), *Experimental approaches to phonology*, 351–368. Oxford: Oxford University Press.

Otake, Takashi. this volume. Chapter 12: Mora and mora timing. In Haruo Kubozono, (ed.), *The handbook of Japanese phonetics and phonology.* Berlin: De Gruyter Mouton.

Otaki, Yoko. 2011. The boundary of the quantitative contrast between singleton/geminate consonants in production and perception. In Wai-Sum Lee and Eric Zee (eds.), *Proceedings of ICPhS XVII*, Hong Kong: City University of Hong Kong, 1546–1549.

Payne, Elinor. 2005. Phonetic variation in Italian consonant gemination. *Journal of International Phonetic Association* 35(2). 153–181.

Payne, Elinor. 2006. Non-durational indices in Italian geminate consonants. *Journal of International Phonetic Association* 36(1). 83–95.

Pickett, Emily, Sheila Blumstein, and Martha Burton. 1999. Effects of speaking rate on the singleton/geminate consonant contrast in Italian. *Phonetica* 56. 135–157.

Pickett, J. M. and Louis R. Decker. 1960. Time factors in the perception of a double consonant. *Language and Speech* 3. 11–17.

Pind, Joergen. 1986. The perception of quantity in Icelandic. *Phonetica* 43. 116–139.

Pintér, Gábor. this volume. Chapter 3: The emergence of new consonant contrasts. In Haruo Kubozono (ed.), *The handbook of Japanese phonetics and phonology.* Berlin: De Gruyter Mouton.

Podesva, Robert. 2000. Constraints on geminates in Burmese and Selayarese. In Roger Bilerey-Mosier and Brook Danielle Lillehaugen (eds.), *Proceedings of West Coast Conference on Formal Linguistics 19*, 343–356. Somerville: Cascadilla Press.

Port, Robert and Johnathan Dalby. 1982. Consonant/vowel ratio as a cue for voicing in English. *Perception* & *Psychophysics* 32. 141–152.

Port, Robert and Michael O'Dell. 1985. Neutralization of syllable-final voicing in German. *Journal of Phonetics* 13. 455–471.

Port, Robert, Jonathan Dalby and Michael O'Dell. 1987. Evidence for mora timing in Japanese. *Journal of the Acoustical Society of America* 81. 1574–1585.

Ridouane, Rachid. 2010. Geminate at the junction of phonetics and phonology. In Cécile Fougeron, Barbara Kühnert, Mariapaola D'Imperio and Nathalie Valleé (eds.), *Laboratory phonology 10*, 61–90. Berlin: Mouton de Gruyter.

Sagisaka, Yoshinori and Yoh'ichi Tohkura. 1984. Kisoku ni yoru onsei gōsei no tame no on'in jikanchō seigyo [Phoneme duration control for speech synthesis by rule]. *Denshi Tsūshin Gakkai Ronbunshi* [The Transactions of the Institute of Electronics, Information and Communication Engineers A] 67(7). 629–636.

Sato, Hirokazu. 1978. Temporal characteristics of spoken words in Japanese. *Journal of the Acoustical Society of America [Abstract]* 64 (S1). 113–114.

Sawashima, Masayuki. 1968. Movements of the larynx in articulation of Japanese consonants. *Annual bulletin research institute of logopedics and phoniatrics* 31(1). 11–20.

Shirai, Setsuko. 2002. Gemination in loans from English to Japanese. MA thesis, University of Washington.

Shrotriya, Nisheeth, A. S. Siva Sarma, Rajesh Verma and S. S. Agrawal. 1995. Acoustic and perceptual characteristics of geminate Hindi stop consonants. In Kjell Elenius and Peter Branderud (eds.), *Proceedings of ICPhS XIII*, 132–135.

Smith, Caroline. 1995. Prosodic patterns in the coordination of vowel and consonant gestures. In Bruce Connell and Amalia Arvaniti (eds.), *Papers in laboratory phonology IV: Phonology and phonetic evidence*, 205–222. Cambridge: Cambridge University Press.

Solé, Maria-Josep. 2002. Aerodynamic characteristics of trills and phonological patterning. *Journal of Phonetics* 30. 655–688.

Stevens, Kenneth and Sheila Blumstein. 1981. The search for invariant acoustic correlates of phonetic features. In Peter D. Eimas and Joanne L. Miller (eds.), *Perspectives on the study of speech*, 1–38. New Jersey: Lawrence Erlbaum.

Stevens, Kenneth and Samuel Jay Keyser. 1989. Primary features and their enhancements in consonants. *Language* 65. 81–106.

Tajima, Keiichi, Hiroaki Kato, Amanda Rothwell, Reiko Akahane-Yamada and Kevin. G. Munhall. 2008. Training English listeners to perceive phonemic length contrasts in Japanese. *Journal of the Acoustical Society of America* 123(1). 397–413.

Takada, Seiji. 1985. Sokuon no chōonjō no tokuchō ni tsuite [Articulatory characteristics of geminates]. In Kokuritsu Kokugo Kenkyūjo (ed.), *Kokuritsu Kokugo Kenkyūjo hōkoku 83: Kenkyūjo hōkokushyū 6*, 17–40. Tokyo: Kokuritsu Kokugo Kenkyūjo.

Takagi, Naoyuki and Virginia Mann. 1994. A perceptual basis for the systematic phonological correspondences between Japanese loan words and their English source words. *Journal of Phonetics* 22. 343–356.

Takeyasu, Hajime. 2009. Masatsuon no sokuon chikaku ni okeru masatsu shūhasū tokusei no eikyō [Effects of the spectral properties of frication on perception of singleton/geminate fricatives]. *On'in Kenkyū* [Phonological Studies] 12. 43–50.

Takeyasu, Hajime. 2012. Sokuon no chikaku ni taisuru senkō onsetsu shiin/boin no jizoku jikan no eikyō [Effects of the consonant and vowel durations in the preceding syllable on the perception of geminate stops in Japanese]. *On'in Kenkyū* [Phonological Studies] 15. 67–78.

Tews, Andrea. 2008. Japanese geminate perception in nonsense words involving German [f] and [x]. *Gengo Kenkyu* 133. 133–145.

Tserdanelis, Georgios and Amalia Arvaniti. 2001. The acoustic characteristics of geminate consonants in Cypriot Greek. In Aggouraki, Yoryia, Amalia Arvaniti, J. I. M. Davy, Dyonysis Goutsos, Marilena Karyolaimou, Anna Panagiotou, Andreas Papapavlou, Pavlos Pavlou and Anna Roussou (eds.), *Proceedings of the Fourth International Conference on Greek Linguistics*, 29–36. Thessaloniki: University Studio Press.

van Dommelen, Wim. 1999. Auditory accounts of temporal factors in the perception of Norwegian disyllables and speech analogs. *Journal of Phonetics* 27. 107–123.

Warner, Natasha and Takayuki Arai. 1999. Japanese mora-timing: A review. *Phonetica* 58. 1–25.

Watanabe, Shinichiro and Nobuo Hirato. 1985. The relation between the perceptual boundary of voiceless plosives and their moraic counterparts and the duration of the preceding vowels. *Onsei Gengo* 1. 1–8.

Westbury, John R. and Patricia Keating. 1986. On the naturalness of stop consonant voicing. *Journal of Linguistics* 22. 145–166.

Whalen, Douglas, Arthur Abramson, Leigh Lisker and Maria Mody. 1993. F0 gives voicing information even with unambiguous voice onset times. *Journal of the Acoustical Society of America* 47. 36–49.

Yu, Alan. 2011. Contrast reduction. In John Goldsmith, Jason Riggle and Alan Yu (eds.), *The handbook of phonological theory*, 2nd edition, 291–318. Oxford: Blackwell.