Durational properties of emphatically lengthened consonants in Japanese*

Shigeto Kawahara and Aaron Braver

Keio University and Texas Tech University

Abstract

Languages can make use of phonetic duration to signal two kinds of meanings. The first is a lexical, phonological contrast. This sort of contrast is usually limited to a binary distinction, and its phonetic properties have been well studied for many different languages. The other use of phonetic duration is to express pragmatic emphasis. Speakers of some languages can use lengthening to express emphasis, as in the English example *Thank you soooooooo much*. This lengthening can employ multiple degrees of duration, beyond the more standard binary contrast. This second use of duration has been understudied, and this paper attempts to fill that gap. To that end, this paper reports the first experimental documentation of the consonant lengthening pattern in Japanese, which expresses pragmatic emphasis. The results show that at least some speakers show six levels of durational distinctions, while other speakers show less clear-cut distinctions among different levels of emphatically lengthened consonants. Nevertheless, all but one speaker showed a linear correlation between duration and level of emphasis.

*The recording for this experiment was conducted while the first author was a visiting scholar at International Christian University, which was made possible by a fellowship offered by the Japan ICU Foundation, for which we are grateful. The help by research assistants at the Rutgers Phonetics Laboratory with the acoustic analysis was indispensable. John Kingston and three anonymous reviewers offered very helpful comments on this project. Remaining errors are ours.

1 Introduction

Languages can make use of phonetic duration to signal two sorts of semantic functions. The first is a lexical, phonological contrast. For example, in Japanese [kata] with a short [t] means 'frame' and [katta] with a long [tt] means 'bought'. In most languages with such a length contrast, the contrast is binary; that is, the distinction is a matter of short vs. long.[1] The phonetic properties of such lexical short-long contrasts have been well studied for many languages; Table 1 summarizes previous phonetic studies on short vs. long consonants in various languages (this list is not exhaustive; see also Ridouane 2010 for another recent summary.).

Less well-studied are cases in which speakers use lengthening to express pragmatic emphasis. For example, in English, speakers can say *Thank you soooooo much* to express an emphatic meaning; in this case, the speaker is trying to express that the speaker's degree of gratitude is very high.[2] An impressionistic observation seems to suggest that this sort of lengthening is not limited to a binary contrast—an intuition which will be confirmed in the experiment reported below. Compared to lexical singleton-geminate distinctions, the phonetic properties of this sort of contrast are understudied in the phonetics literature, and our experiment aims to fill this gap.

As a case study, we investigate the durational properties of emphatically lengthened consonants found in casual speech of Japanese, in which lengthen segments to express emphasis (Aizawa, 1985; Kawahara, 2001, 2013; Nasu, 1999).[3] This emphatic lengthening is a characteristic of casual speech by young speakers, and frequently appears (orthographically) on the internet and in comic

[1] There are rare cases in which the contrast is arguably ternary: short vs. long vs. overlong (e.g., Estonian: Prince 1980, and Saami: Bals Baal et al. 2012). However, ternary length contrasts are very rare cross-linguistically (Ladefoged and Maddieson, 1996, p. 93), and even in languages that do have ternary length contrasts, morphological and other factors are likely to affect the distribution of overlong consonants. As Prince puts it (1980: 511), in Estonian "the three-way contrast, and in particular the distribution of overlength, is richly and curiously connected with patterns of morphology, syllable structure, and stress." See also Bals Baal et al. (2012) for similar complications related to the case of the ternery consonant length contrast in Northern Saami.

[2] The terms "emphasis" and "emphatic" used in this paper are different from so-called "emphatic consonants" found in Arabic and other languages (Bellem, 2007; Kahn, 1975; Laufer and Baer, 1988; McCarthy, 1994; Norlin, 1987). The crucial difference is that the emphatic consonants in Arabic are separate phonemes, opposed to non-emphatic counterparts. What the current experiment deals with is not a lexical contrast, but lengthening due to pragmatic emphasis.

[3] Japanese speakers can also lengthen vowels to express this sort of pragmatic emphasis. See Kawahara and Braver (2013) for the phonetic properties of emphatically lengthened vowels. There is no overlap between the participants of the current experiment and those of Kawahara and Braver (2013).

Table 1: A summary of phonetic studies on lexical duration-based contrasts in consonants.

Language	**References**
Arabic	Ham (2001); Norlin (1987)
Bengali	Lahiri and Hankamer (1988)
Berber	Ridouane (2010)
Bernese	Ham (2001)
Buginese	Cohn et al. (1999)
Chicasaw	Gordon et al. (2000)
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)
Icelandic	Games (1976)
Italian	Esposito and Di Benedetto (1999); Lisker (1958); Payne (2005); Pickett et al. (1999)
Itunyoso Trique	DiCanio (2012)
Japanese	Han (1962); Homma (1981); Idemaru and Guion (2008); Kawahara (to appear)
Jawon	Jaeger (1983)
Madurese	Cohn et al. (1999); Ham (2001)
Malayalam	Local and Simpson (1999)
Marathi	Lisker (1958)
Pattani Malay	Abramson (1987)
Persian	Hansen (2004)
Rembarrnga	McKay (1980)
(Lule) Saami	Engstrand (1987)
Swedish	Engstrand and Krull (1994); Lisker (1958)
Swiss German	Krähenmann (2003); Kraehenmann and Lahiri (2008)
Toba Batak	Cohn et al. (1999)
Turkish	Lahiri and Hankamer (1988)
Zapotec	Jaeger (1983)

books. In this phenomenon, there can be multiple degrees of durational differences, beyond the standard short-long binary distinction. In Japanese, gemination is expressed orthographically with a small diacritic symbol (っ) preceding the mora containing the consonant in question, as shown in example (b) in Table 2. The emphaitic lengthening that is at issue here can be expressed by the use of the same gemination marker. For example, Japanese speakers can take an adjective like (c), and geminate the (word-medial) consonant to express emphatic meaning, as in (d). In casual writing, we observe examples in which consonants are accompanied by a number of gemination marks, as in (e-h).

Table 2: The Japanese orthographic system for gemination. The forms in (a) and (b) represent a lexical singleton-geminate pair. The forms in (d)-(h) represent emphatically lengthened geminates, which are the focus of investigation in the current study.

Japanese orthography	**Transcription**	**Gloss**
a. かた	[kata]	'frame'
b. かった	[katta]	'bought'
c. かたい	[katai]	'hard'
d. かったい	[kattai]	'hard' (emphatic)
e. かっったい	[katttai]	'hard' (emphatic)
f. かっっったい	[kattttai]	'hard' (emphatic)
g. かっっっったい	[katttttai]	'hard' (emphatic)
h. かっっっっったい	[kattttttai]	'hard' (emphatic)

The aim of this project is to investigate the durational characteristics of this multiple-level emphasis pattern, the primary question being how many levels of distinctions speakers can actually realize acoustically in this sort of pragmatically-driven lengthening. While the phonetic properties of Japanese lexical geminates have been investigated in many instrumental studies in the past (see Kawahara to appear for a recent overview), the current multiple emphasis pattern has not been investigated from a phonetic/instrumental perspective. This paper thus offers the first experimental documentation of this multiple emphasis pattern. More generally speaking, the phonetics of pragmatically lengthened segments has been less well studied than the phonetics of lexical short-long contrasts, and our study aims to provide extensive documentation of the first kind of lengthening.

2 Method

2.1 Stimuli

This study measured the duration of four coronal obstruents, [t, d, s, z], as used in the emphatic environments.[4] For each sound, two adjectives were chosen, since adjectives are (semantically speaking) most likely to undergo emphasis. The adjectives used in this experiment, listed in Table 3, were all disyllabic and lexically accented on the second syllable (i.e., they all had an HL falling pitch contour on the second syllable). The target consonants were always placed in word-medial position. Each adjective was paired with a subject noun phrase to make a complete sentence: e.g. [ano koogi uzai] 'that lecture is annoying'.

Table 3: List of stimuli. Two adjectives for each consonant were chosen.

[t]	[d]	[s]	[z]
katai 'hard'	hidoi 'awful'	kusai 'smelly'	uzai 'annoying'
itai 'aching'	kudoi 'wordy'	musai 'disgusting'	mazui 'distasteful'

For each adjective, in addition to a non-emphatic rendition, five degrees of emphasis were created; e.g., [katai] (no emphasis), [kattai] (level 1 emphasis), [katttai] (level 2), [kattttai] (level 3), [katttttai] (level 4), and [kattttttai] (level 5), as illustrated in examples (c-h) in Table 2.

As a result, there were a total of 48 stimuli (4 consonants $\times$ 2 adjectives $\times$ 6 consonant lengths). A random number was assigned to each stimulus item to track which stimulus was actually pronounced.

[4] Japanese does not possess approximant geminates (Kawahara, to appear). Japanese does have nasal geminates, but geminating nasals for emphatic purposes is disfavored (Kawahara, 2013). This study therefore focused on obstruent geminates. In order to control for the effect of place of articulation on duration (Maddieson, 1997), the experiment used only coronal consonants.

2.2 Participants

The participants were seven native speakers of Japanese (Speakers FR, FV, SX, EG, NN, LV, TV). They were all undergraduate students at International Christian University (Tokyo, Japan). They were paid 500 Japanese yen for their time. They were all in their twenties at the time of recording.

2.3 Procedure

The experiment took place in a sound-attenuated room at International Christian University. Superlab version 4.0 was used to present the stimuli (Cedrus Corporation, 2010). The stimuli and the instructions were all presented in Japanese orthography. In the instructions, speakers were told that the experiment was about multiple levels of emphasis in Japanese, i.e., that they were going to be reading sentences with multiple gemination marks.

In order to prevent them from resorting to explicit counting of gemination marks by way of gesture, they were asked not to use gestures such as counting using fingers or nodding their heads. They were also told that their goal was not to count the gemination marks, but rather to pronounce Japanese utterances that are suitable for each stimulus. They were also told that the experiment was not a competition or a test, and that the experiment was instead a test of what Japanese speakers actually do (i.e., the experiment was not for a prescriptive, but rather a descriptive purpose).

Each block contained all the stimuli. After each block, the speaker took a short break. The order within each block was randomized by Superlab within each repetition. The speakers were asked to go through eight blocks (48×8=384 tokens), although due to time limitations, one speaker (Speaker NN) could only complete six repetitions (each speaker was assigned 30 minutes for this experiment because of a scheduling restriction). Some speakers mispronounced or skipped a few tokens.

As a practice session before the main session, all the speakers went through all the stimuli once to familiarize themselves with the stimuli and the task. After the practice phase, they were allowed to ask any questions that they had.

Their pronunciation was directly recorded into a portable recorder (TASCAM DR-40) with a

44.1k sampling rate and a 16 bit quantization level. The experimenter (the first author) sat with the speakers throughout the experiment.

2.4 Acoustic analysis

To investigate the acoustic realizations of multiple-emphasis patterns, this study focused on consonant durations, because they are the main acoustic correlate of Japanese length contrasts (Amano and Hirata, 2010; Beckman, 1982; Han, 1962, 1992, 1994; Hirata and Whiton, 2005; Hirata and Amano, 2012; Hirose and Ashby, 2007; Homma, 1981; Idemaru and Guion, 2008; Kawahara, 2006). There are other acoustic covariants of gemination in Japanese (Idemaru and Guion, 2008; Kawahara, 2006), and a post-hoc analysis on preceding vowel duration is reported in section 3.3.

The boundary between the target consonants and the surrounding vowels was placed by inspecting both the waveforms (onset and offset of aperiodic noise for the fricatives and stop closure for stops) and spectrograms (abrupt cessation of F2 and F3 in particular). Figures 1 and 2 illustrate sample waveforms and spectrograms of three tokens of [t] and [s] (no emphasis, level 1 emphasis, level 2 emphasis)—the time scales are all 1,000ms. The acoustic analysis was performed using Praat (Boersma, 2001; Boersma and Weenink, 1999–2013).

2.5 Statistics

Since there are many comparisons (6 levels of emphasis $\times$ 4 types of consonants $\times$ 2 adjectives = 48 comparisons for each speaker), to avoid Type I error, we did not conduct pair-wise comparisons of every condition. Instead, we compared each level of emphasis by collapsing the consonant types and adjective types, thereby making only 5 pair-wise comparisons for each speaker. By Bonferroni adjustment, the α-level was set to be $0.05/5 = 0.01$.[5] In addition, post-hoc inspection of the data also suggested that regression analyses would be useful, so they are reported in the results section.

[5]We thank an anonymous reviewer for useful advice on this point. To provide a measure of how much variability exists in the current data, error bars, which represent 95% confidence intervals, are also provided in the result figures. They were generally calculated over 16 repetitions of each consonant (2 adjectives $\times$ 8 repetitions), except for Speaker NN, who pronounced the stimuli 6 times each (see above).

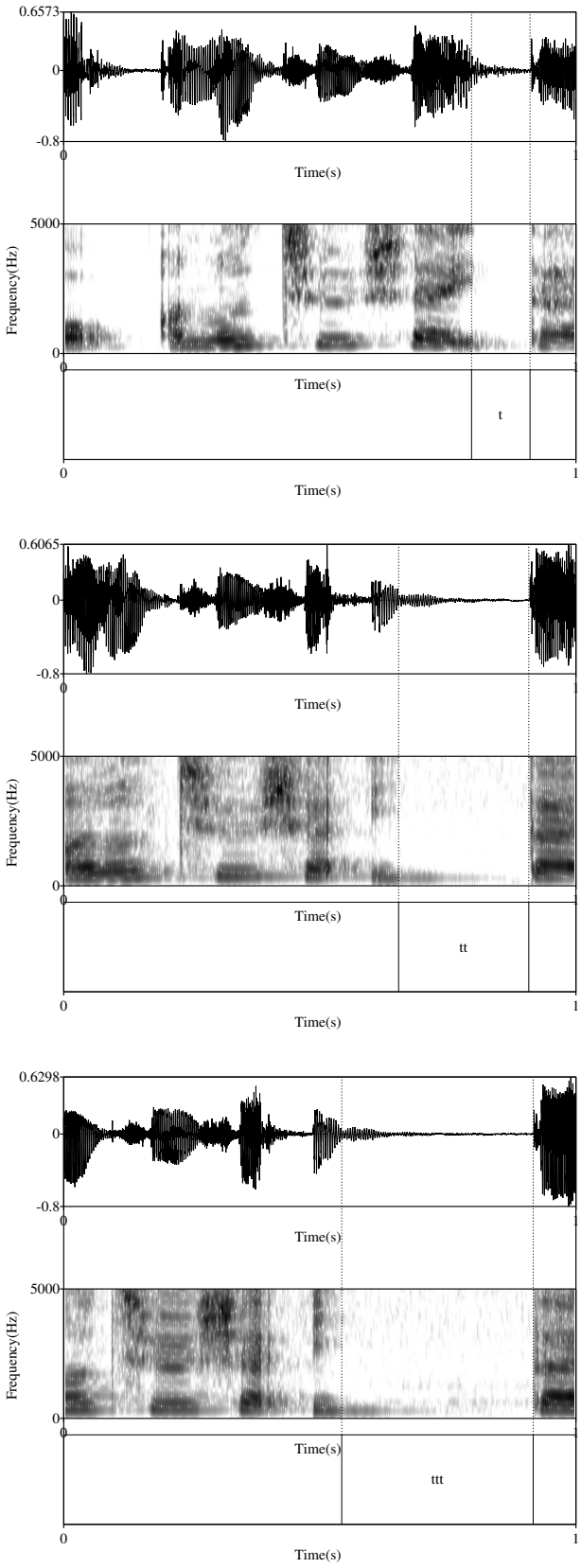

Figure 1: Sample waveforms and spectrograms of [t], pronounced in [it(tt)ai]. The time scales are all 1,000ms.

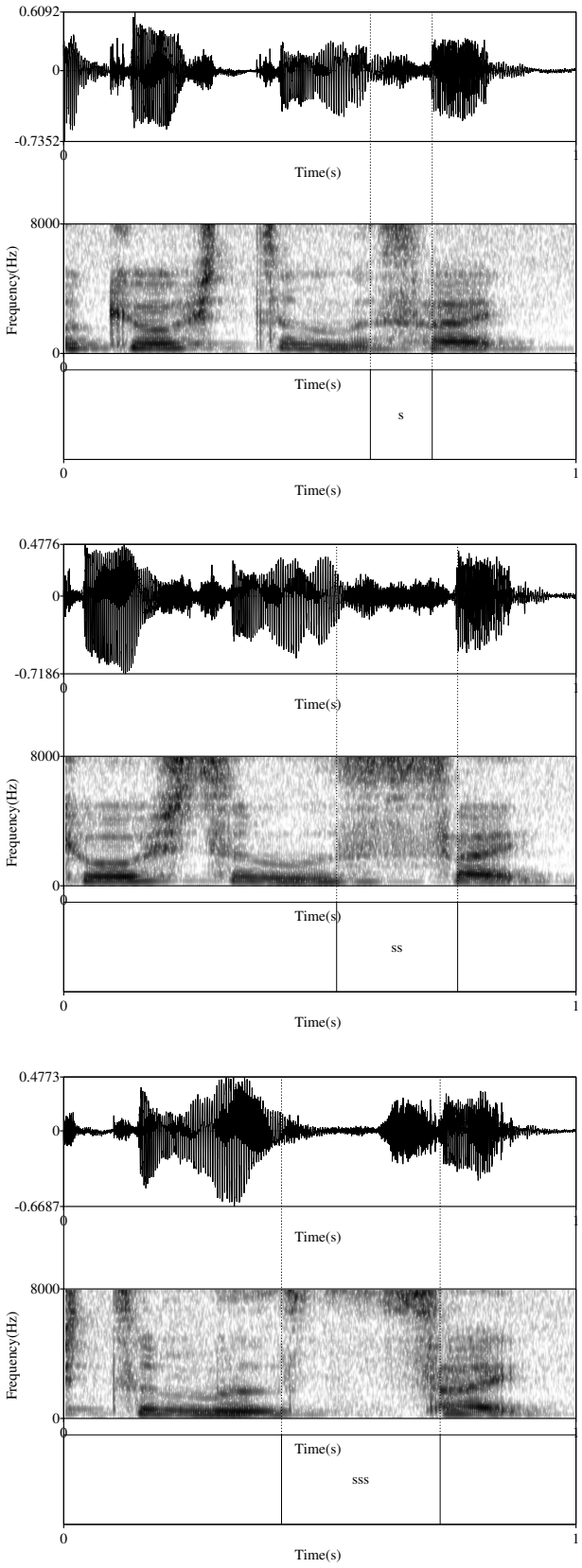

Figure 2: Sample waveforms and spectrograms of [s], pronounced in [mus(ss)ai]. The time scales are all 1,000ms.

All statistical analyses were performed using R (R Development Core Team, 1993–2013).

3 Results

3.1 Individual patterns

Since inter-speaker differences were apparent in the results, the results of individual speakers are reported separately. We discuss each speaker in order of how clearly they showed durational differences among different emphasis levels. The result figures have different y-axis scales, as different speakers used different durational ranges. After examining the behavior of each speaker, we summarize and compare the behaviors of all speakers in section 3.2.

First, of the seven speakers, two speakers (Speakers FR and TW) seem to make a perfect six way distinction; i.e., the consonant durations for each level of emphasis are different. The results of these speakers are illustrated in Figures 3 and 4.

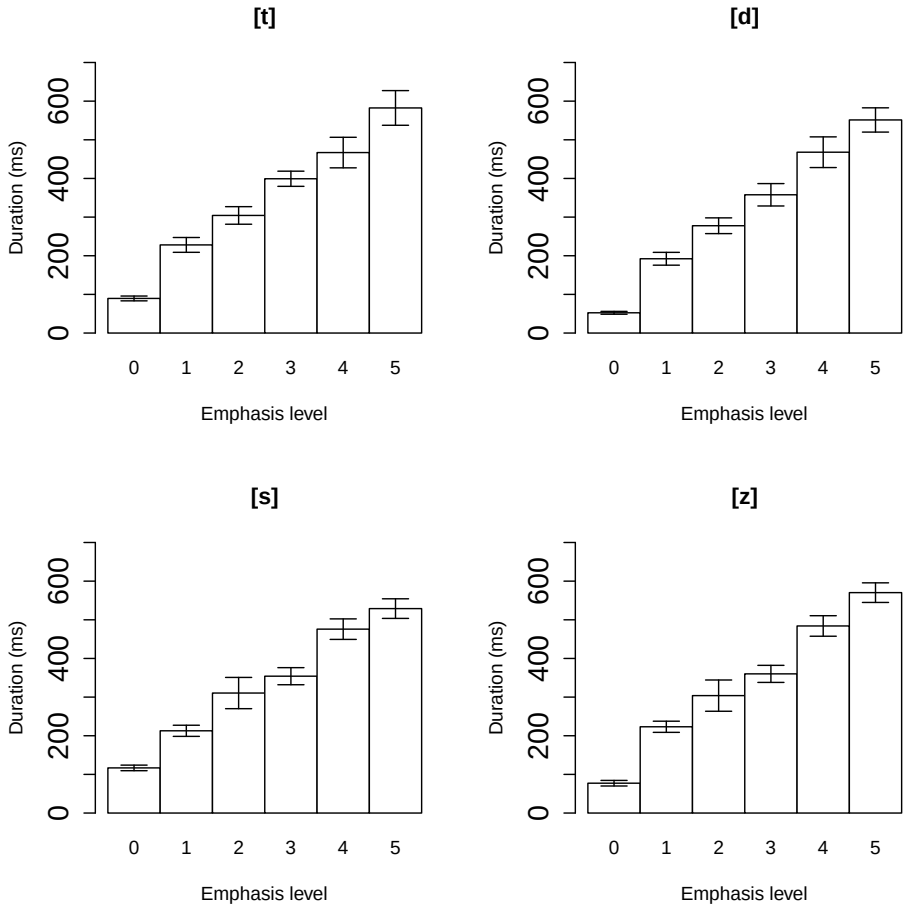

Figure 3: The average durations of each emphasis level with 95% confidence intervals: Speaker FR.

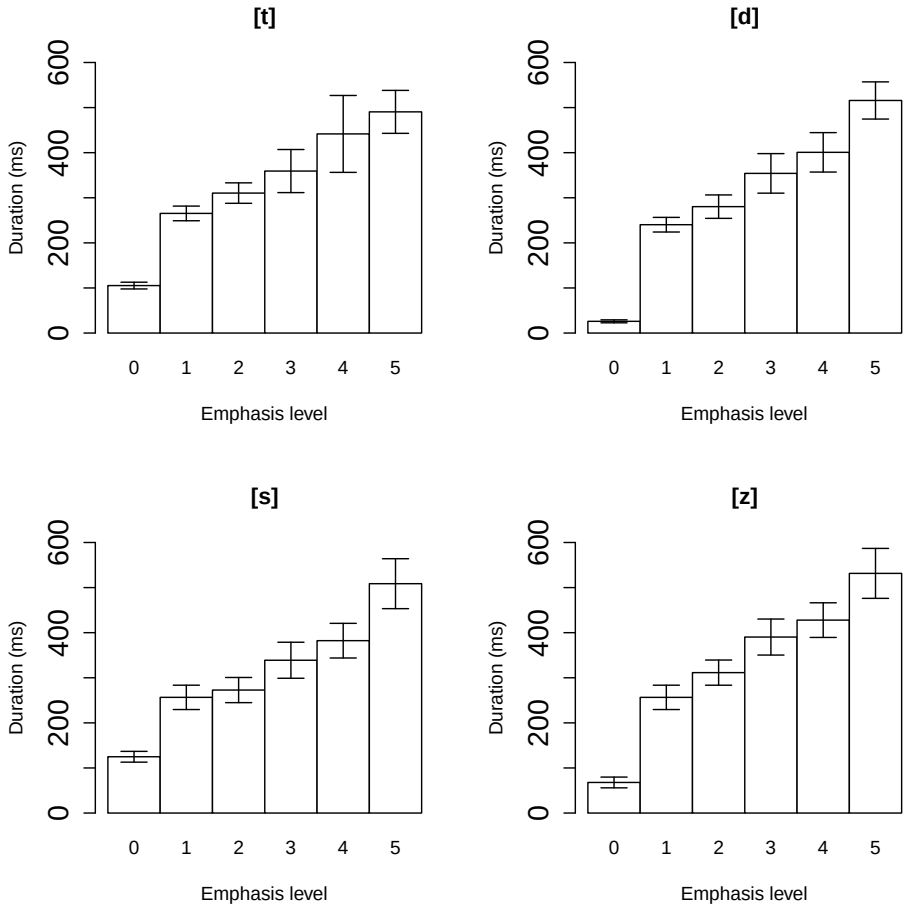

Figure 4: The average durations of each emphasis level: Speaker TW.

We observe that for both speakers there is a large increase in duration from plain consonants to the level 1 emphatically lengthened consonants. Further, within the emphatically lengthened consonants there is a steady, linear increase in duration as the emphasis level increases. Non-paired multiple comparison t-tests show that all levels are different in terms of their duration. These results are given in Table 4.

Table 4: Non-paired multiple comparison t-tests for Speakers FR and TW, showing the effect of emphasis level on duration. α=.01 after Bonferroni adjustment (.05/5).

Speaker FR			**Speaker TW**		
Comparison	***t(df)***	***p***	**Comparison**	***t(df)***	***p***
level 0 vs. level 1	$t(126) = 24.9$	$p < .001$	level 0 vs. level 1	$t(126) = 25.5$	$p < .001$
level 1 vs. level 2	$t(126) = 11.1$	$p < .001$	level 1 vs. level 2	$t(126) = 5.2$	$p < .001$
level 2 vs. level 3	$t(126) = 7.6$	$p < .001$	level 2 vs. level 3	$t(125) = 5.4$	$p < .001$
level 3 vs. level 4	$t(125) = 9.5$	$p < .001$	level 3 vs. level 4	$t(125) = 3.0$	$p < .01$
level 4 vs. level 5	$t(125) = 6.7$	$p < .001$	level 4 vs. level 5	$t(125) = 5.4$	$p < .001$

In addition, to assess the linear correlation between emphasis levels and duration within different levels of emphatically lengthened consonants, a linear regression was run with duration as the dependent variable and with emphasis level as the independent variable (no-emphasis consonants were not included in this regression analysis because of the non-linearity we observe between no-emphasis consonants and emphatically lengthened consonants). For both speakers, the effect of the emphasis level is significant ($t(317) = 38.0$, $p < .001$ for Speaker FR and $t(315) = 19.6$, $p < .001$ for Speaker TW). The estimated coefficients of emphasis level are 86ms and 63ms, respectively—these values are estimates of how many milliseconds these speakers increase a consonant's duration per emphasis level.

Finally, to numerically assess the strength of the correlation between emphasis levels and duration, Pearson correlation coefficients (r) were calculated. The no-emphasis consonants were excluded from this analysis also, because there are large jumps in duration between plain consonants and the emphatically lengthened consonants. The results show that r values are .91 for Speaker FR and .74 for Speaker TW, both very high correlations (both significant at the $p < .001$ level).

Other speakers also showed a steady increase in duration, but not as clearly as Speakers FR and TW. Speaker EL shows the next highest correlation between emphasis level and duration, as shown in Figure 5.

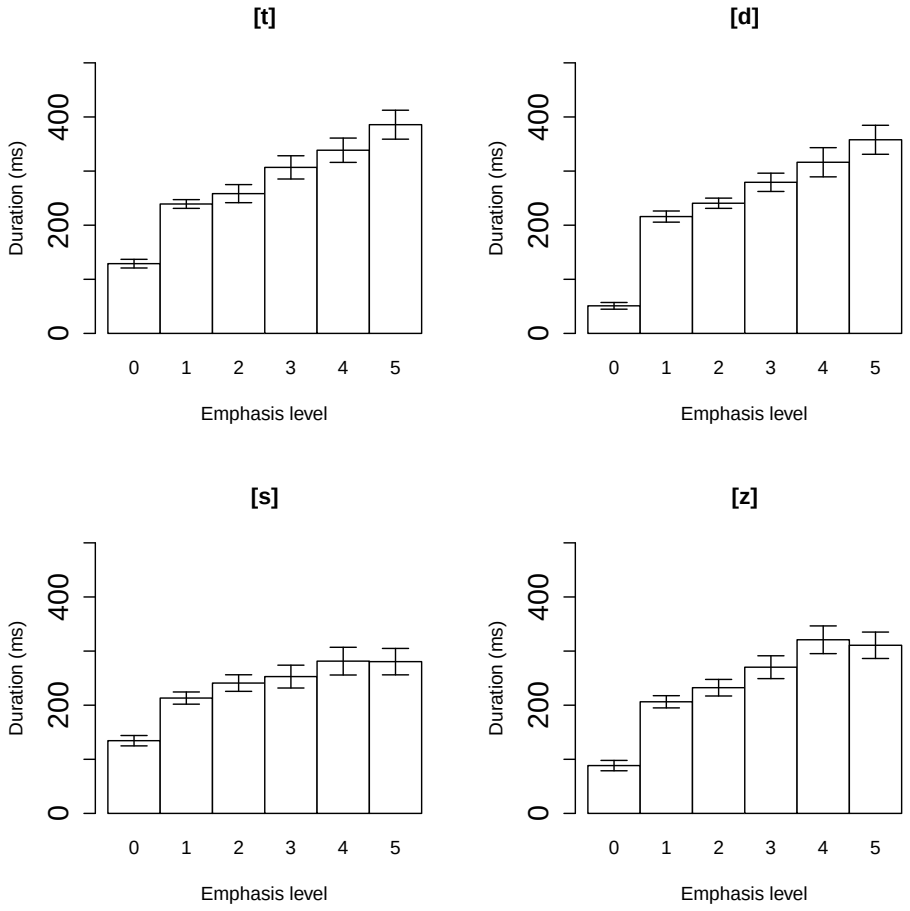

Figure 5: The average durations of each emphasis level: Speaker EL.

Although this speaker does not show a difference between level 4 and level 5 for the two fricatives, there seems to be a clear, general increase of duration as the emphasis levels go higher. The results of multiple comparison t-tests, given in Table 5, show that all the differences but the comparison between level 4 and level 5 are significant.

Table 5: Non-paired multiple comparison t-tests for Speaker EL, showing the effect of emphasis level on duration. α=.01 after Bonferroni adjustment (.05/5).

Speaker EL		
Comparison	***t(df)***	***p***
level 0 vs. level 1	$t(126) = 21.5$	$p < .001$
level 1 vs. level 2	$t(126) = 5.5$	$p < .001$
level 2 vs. level 3	$t(126) = 5.5$	$p < .001$
level 3 vs. level 4	$t(125) = 4.5$	$p < .001$
level 4 vs. level 5	$t(125) = 1.9$	*n.s.* ($p = 0.06$)[6]

The effect of emphasis is statistically significant in the regression model ($t(317) = 17.3, p < .001$), and the coefficient estimate is 30ms. Despite the fact that this speaker does not show differences for some levels of emphasis, the r-value for Speaker EL is high ($r = .70, p < .001$). We also notice that the duration range is smaller (about 500ms in Figure 5) compared to the previous two speakers (about 700ms and 600ms in Figures 3 and 4, respectively), and thus this speaker manages to—or at least attempts to—make six levels of duration distinctions within a smaller duration range. This characteristic is perhaps responsible for the smaller estimate of the effect of emphasis in the regression model.

The next speaker, Speaker SX, shows some increase in duration correlating with emphasis levels, but we observe a number of emphasis pairs that are not differentiated from one another, as shown in Figure 6.

[6] A post hoc test comparing only stops shows that the difference between level 4 and level 5 is significant ($t(62) = 3.6, p < .001$).

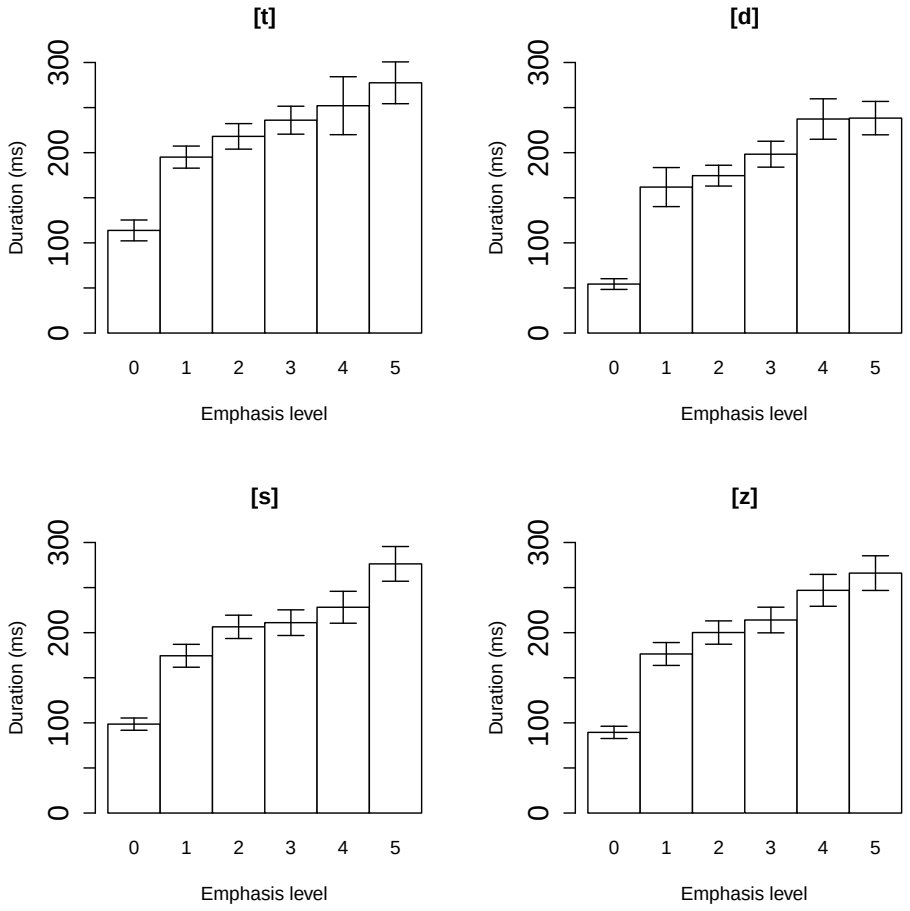

Figure 6: The average durations of each emphasis level: Speaker SX.

The speaker does not show a difference from level 2 to level 4 for [s], or between level 1 and level 2, or level 4 and level 5, for [d]. We also note that this speaker's duration range is even smaller than that of Speaker EL (the maximum range is about 300ms in Figure 6).

The effect of emphasis is still significant in the regression model ($t(314) = 15.5, p < .001$), but the coefficient estimate is lower (22ms), compared to the previous two speakers. The smaller coefficient is presumably related to the fact that the duration range is smallest among the speakers we have seen thus far. This speaker's r-value is slightly lower than the previous speakers' r-values ($r = .66, p < .001$). Nevertheless, all the differences turned out to be significant by multiple comparison t-tests, the results of which are given in Table 6, presumably because each level of difference is manifested by some if not all consonants.

Table 6: Non-paired multiple comparison t-tests for Speaker SX, showing the effect of emphasis level on duration. α=.01 after Bonferroni adjustment (.05/5).

Speaker SX		
Comparison	***t(df)***	***p***
level 0 vs. level 1	$t(126) = 17.6$	$p < .001$
level 1 vs. level 2	$t(126) = 4.3$	$p < .001$
level 2 vs. level 3	$t(126) = 2.8$	$p < .01$
level 3 vs. level 4	$t(122) = 4.1$	$p < .001$
level 4 vs. level 5	$t(122) = 3.2$	$p < .001$

Next, as shown in Figure 7, Speaker EG often fails to show differences between emphasis levels in the middle range (between level 1 and level 2 as well as between level 3 and level 4 for [s], and from level 1 to level 3 for the two voiced consonants). The statistical tests, given in Table 7, show no significant differences between level 1 and level 3, but significant results elsewhere.

It seems that this speaker has a four way contrast: non-emphatic (level 0), emphatic (level 1 to 3), very emphatic (level 4), and most emphatic (level 5). As an anonymous reviewer pointed out, this speaker may only have this four-way internal representations, and translated the multiple degrees of gemination marks to fit into these categories.

The effect of emphasis is nevertheless significant in the regression analysis ($t(310) = 15.1, p < .001$), and the coefficient estimate is higher than Speaker SX (52ms). Despite the apparent lack of

Table 7: Non-paired multiple comparison t-tests for Speaker EG, showing the effect of emphasis level on duration. α=.01 after Bonferroni adjustment (.05/5).

Speaker EG		
Comparison	***t(df)***	***p***
level 0 vs. level 1	$t(125) = 26.3$	$p < .001$
level 1 vs. level 2	$t(124) = 2.5$	*n.s.* ($p = .012$)
level 2 vs. level 3	$t(124) = 1.8$	*n.s.* ($p = .07$)
level 3 vs. level 4	$t(121) = 4.2$	$p < .001$
level 4 vs. level 5	$t(121) = 6.4$	$p < .001$

differences in the middle range, r is reasonably high ($r = .65$, $p < .001$).

Speaker FV, shown in Figure 8, does show a steady increase in duration, but we observe that the speaker does not show a difference between certain emphasis levels; e.g. level 1 and level 2 as well as level 3 and level 4 for [t]; level 4 and level 5 for [d]; level 3 and level 4 for [s]; level 1 to level 3, and level 4 to level 5 for [z]. The statistical tests show that, after Bonferroni correction, only the difference between level 0 and level 1 and the difference between level 2 and level 3 are reliable, as can be seen in Table 8. The effect of emphasis is still significant in the regression analysis ($t(275) = 11.3$, $p < .001$), but the coefficient estimate is low (24ms). r is also low ($r = .56$, $p < .001$), as compared to the other speakers we have seen.

Table 8: Non-paired multiple comparison t-tests for Speaker FV, showing the effect of emphasis level on duration. α=.01 after Bonferroni adjustment (.05/5).

Speaker FV		
Comparison	***t(df)***	***p***
level 0 vs. level 1	$t(108) = 13.7$	$p < .001$
level 1 vs. level 2	$t(110) = 2.4$	*n.s.* ($p = ..018$)
level 2 vs. level 3	$t(110) = 2.6$	$p < .01$
level 3 vs. level 4	$t(109) = 2.3$	*n.s.* ($p = .02$)
level 4 vs. level 5	$t(109) = 2.5$	*n.s.* ($p = .012$)

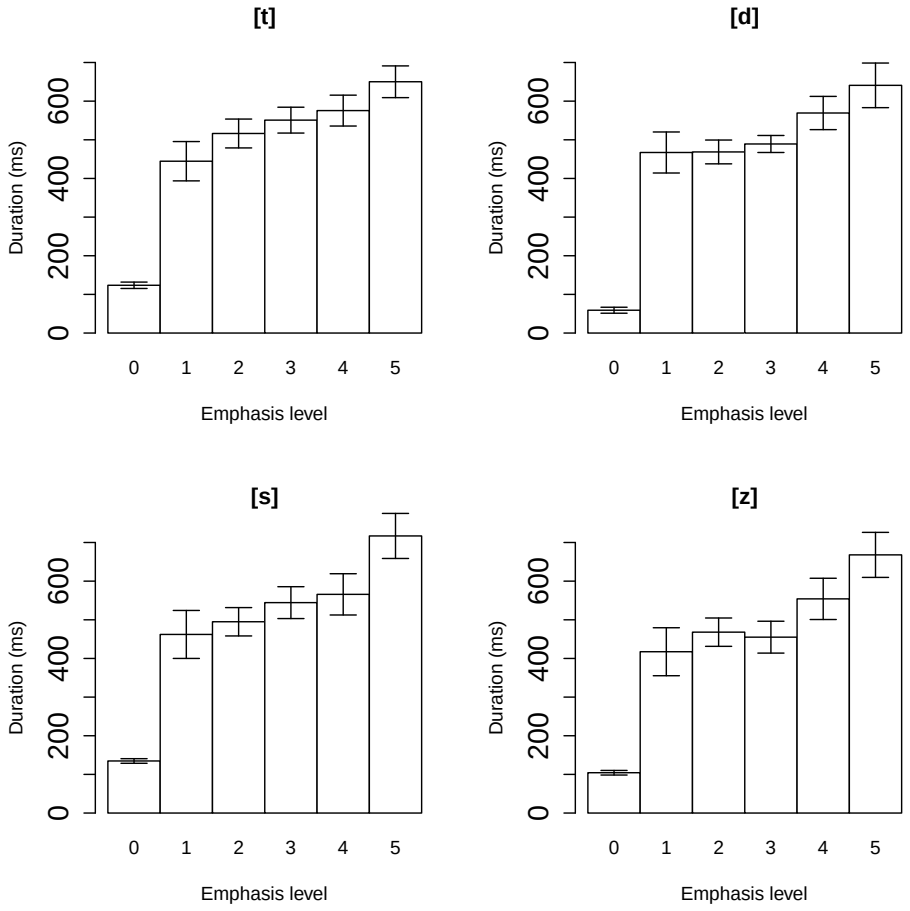

Figure 7: The average durations of each emphasis level: Speaker EG.

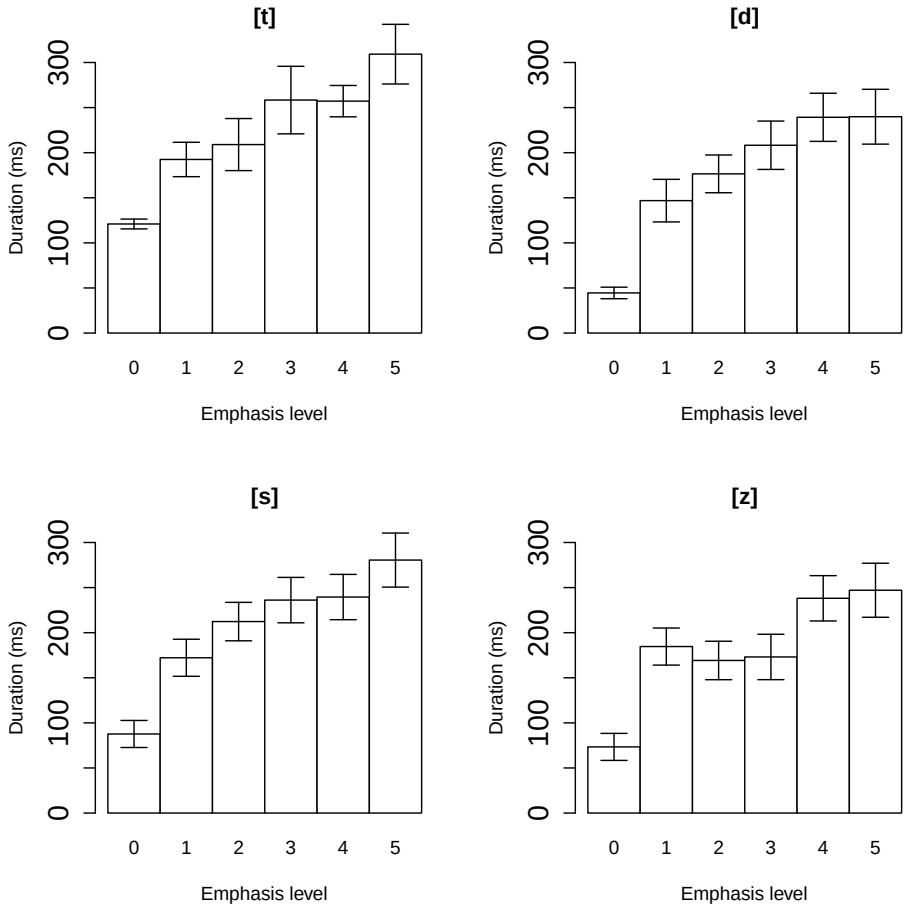

Figure 8: The average durations of each emphasis level: Speaker FV.

Finally, one speaker (Speaker NN) showed a more or less binary distinction—i.e., plain vs. emphatically lengthened, as shown in Figure 9. That is, this speaker does not seem to show distinctions among different levels of emphatically lengthened consonants (and this speaker seems to show an evident reversal between level 1 and level 2 for [z]). The results of multiple comparison t-tests, given in Table 9, support this observation. Although statistically significant ($t(236) = 3.33, p < .001$), the coefficient estimate in the regression model is as small as 7ms. The r value is also very low ($r = .21, p < .001$), compared to the other speakers. In short, this speaker may allow only two categories—no-emphasis and emphatically-lengthened—without any further distinctions among emphatically lengthened consonants.

Table 9: Non-paired multiple comparison t-tests for Speaker NN, showing the effect of emphasis level on duration. α=.01 after Bonferroni adjustment (.05/5).

Speaker NN		
Comparison	***t(df)***	***p***
level 0 vs. level 1	$t(93) = 10.5$	$p < .001$
level 1 vs. level 2	$t(92) = 0.0$	*n.s.* ($p = .97$)
level 2 vs. level 3	$t(93) = 2.6$	*n.s.* ($p = .011$)
level 3 vs. level 4	$t(94) = 0.3$	*n.s.* ($p = .75$)
level 4 vs. level 5	$t(94) = 0.2$	*n.s.* ($p = .85$)

3.2 Summary of the patterns of closure duration

Table 10 provides a summary of each speaker's behavior. It provides the regression function for each speaker, as well as the r value. The coefficients represent how many milliseconds each speaker increases consonant duration per emphasis level.[7] The r values are a measure of the strength of the linear correlation between emphasis levels and duration. In addition, as a measure of their duration

[7]Fujisaki et al. (1975) showed that the just noticeable difference (jnd) in duration for Japanese listeners is about 10ms for non-speech pure tones whose base duration was 100ms. The discrimination of durational differences is affected by various factors, including base duration (Abel, 1972; Kato et al., 2002), the spectral nature of the intervals under question (Kato et al., 2002) and intensity changes from the surrounding intervals (Kato and Tsuzaki, 1994; Kato et al., 1997; Kawahara, 2012). Therefore, whether the durational differences exhibited by speakers are perceptible or not must be tested in a separate perception experiment, although the coefficients are all larger than 10ms (with the exception of Speaker NN).

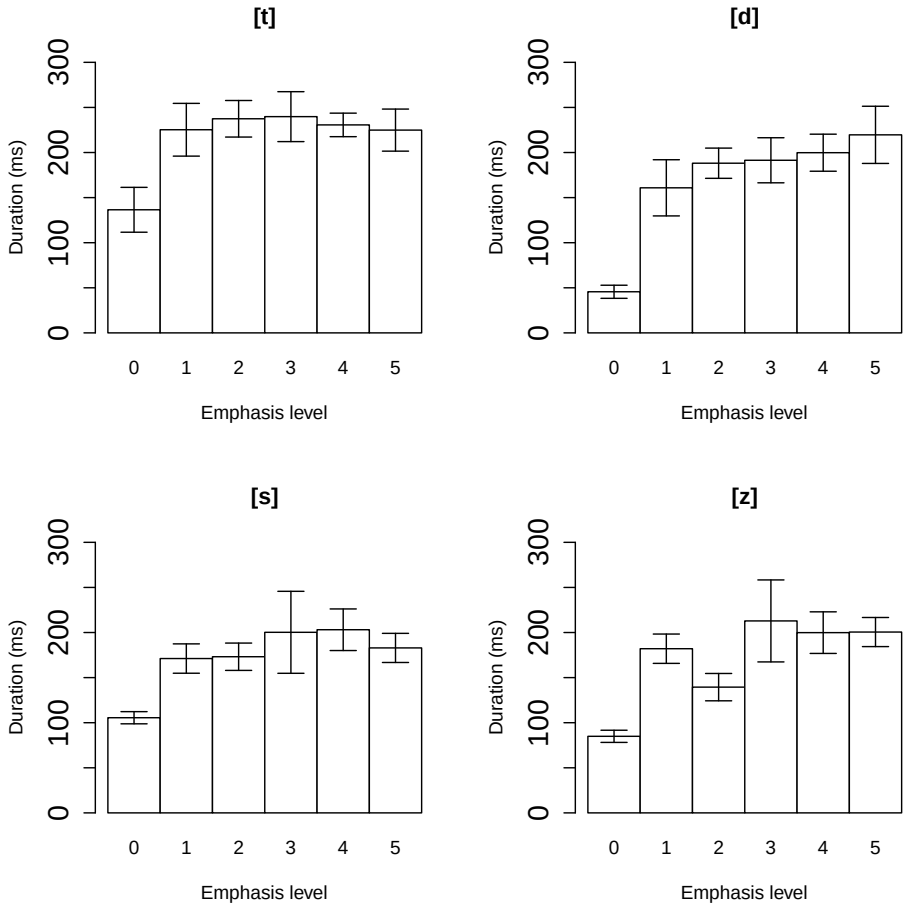

Figure 9: The average durations of each emphasis level: Speaker NN.

range, the maximum duration is provided for each speaker.

All speakers showed a positive correlation between duration and emphasis levels. There are only a few instances of evident reversals, although in a number of cases we observed no differences between certain levels of emphasis.

We also observed that there are noticeable differences among speakers. Two speakers (FR and TW) showed perfect six way distinctions. One speaker (EL) showed some cases in which no differences were observed, but the correlation between emphasis levels and duration was still high. Speaker EG seems to have had four categories. One speaker (NN) made little distinction among emphatically lengthened consonants, although there was a very weak correlation between duration and emphasis levels.

Table 10: Summary of each speaker's behavior

Speaker	**Regression function**	r	**Max Duration (Range)**
Speaker FR	$y = 124 + 86x$	.91	748
Speaker TW	$y = 177 + 63x$	.74	804
Speaker EL	$y = 187 + 30x$	.70	456
Speaker SX	$y = 155 + 22x$	.66	371
Speaker EG	$y = 379 + 52x$	.65	888
Speaker FV	$y = 146 + 24x$	.56	453
Speaker NN	$y = 179 + 7x$	.21	399

In Table 10, we observe an association between how finely each speaker realizes different degrees of emphasis and duration range: for example, Speakers FR and TW, who showed a fine six way distinction, have very large duration ranges. Speaker NN, who showed an almost binary contrast between plain consonants and emphatically lengthened consonants, has a small duration range. The correlation is not perfect, however, since for example, Speaker EL has a high r-value but nevertheless has a relatively small duration range.

To summarize, all speakers showed a positive correlation between emphasis level and consonant duration, although we also observe some inter-speaker variability. Some speakers (especially Speakers FR and TW) seem to have managed to perfectly distinguish six levels of consonantal duration differences. The current experiment included (only) up to level 5 emphasis; it remains to be seen where the limit lies with respect to how many levels of emphasis can actually be produced.

3.3 Is the effect of lengthening localized? The effect on the preceding vowels

This study focused on consonant duration, because the main acoustic correlate of Japanese geminates is constriction duration (Amano and Hirata, 2010; Beckman, 1982; Han, 1962, 1992, 1994; Hirata and Whiton, 2005; Hirata and Amano, 2012; Hirose and Ashby, 2007; Homma, 1981; Idemaru and Guion, 2008; Kawahara, 2006). However, in addition to the results given above, a question arises as to whether, when Japanese speakers are expressing emphasis, the effect of emphasis

is localized to only the target consonants. To address this question, a post-hoc analysis examined the duration of preceding vowels using a subset of the data (namely, instances of the stimulus item [katai] 'hard', as this word provided the best environment for duration measurement of preceding vowels among our existing stimuli).[8]

The results are shown in Figure 10 for each speaker (ordered as per the discussion in Section 3.1 and Table 10). We observe, first of all, that all speakers show longer preceding vowels in the emphatically lengthened condition than in the plain consonant conditions. This observation matches well with the earlier observation about Japanese that preceding vowels are longer before geminates than before singletons (Campbell, 1999; Fukui, 1978; Han, 1994; Hirata, 2007; Hirose and Ashby, 2007; Idemaru and Guion, 2008; Kawahara, 2006; Ofuka, 2003; Port et al., 1987; Takeyasu, 2012). This difference in duration between those preceding plain consonants and those preceding emphasized consonants, shown in Figure 10, is thus as expected from what we know about Japanese lexical geminates.

[8]Although the durations of following vowels are also known to be affected by the singleton-geminate difference in Japanese (Campbell, 1999; Han, 1994; Hirata, 2007; Idemaru and Guion, 2008; Ofuka, 2003), they were not analyzed here because previous studies show that the influence of geminates is smaller on following vowels both acoustically and perceptually (Hirata, 2007; Hirato and Watanabe, 1987; Idemaru and Guion-Anderson, 2010; Ofuka et al., 2005).

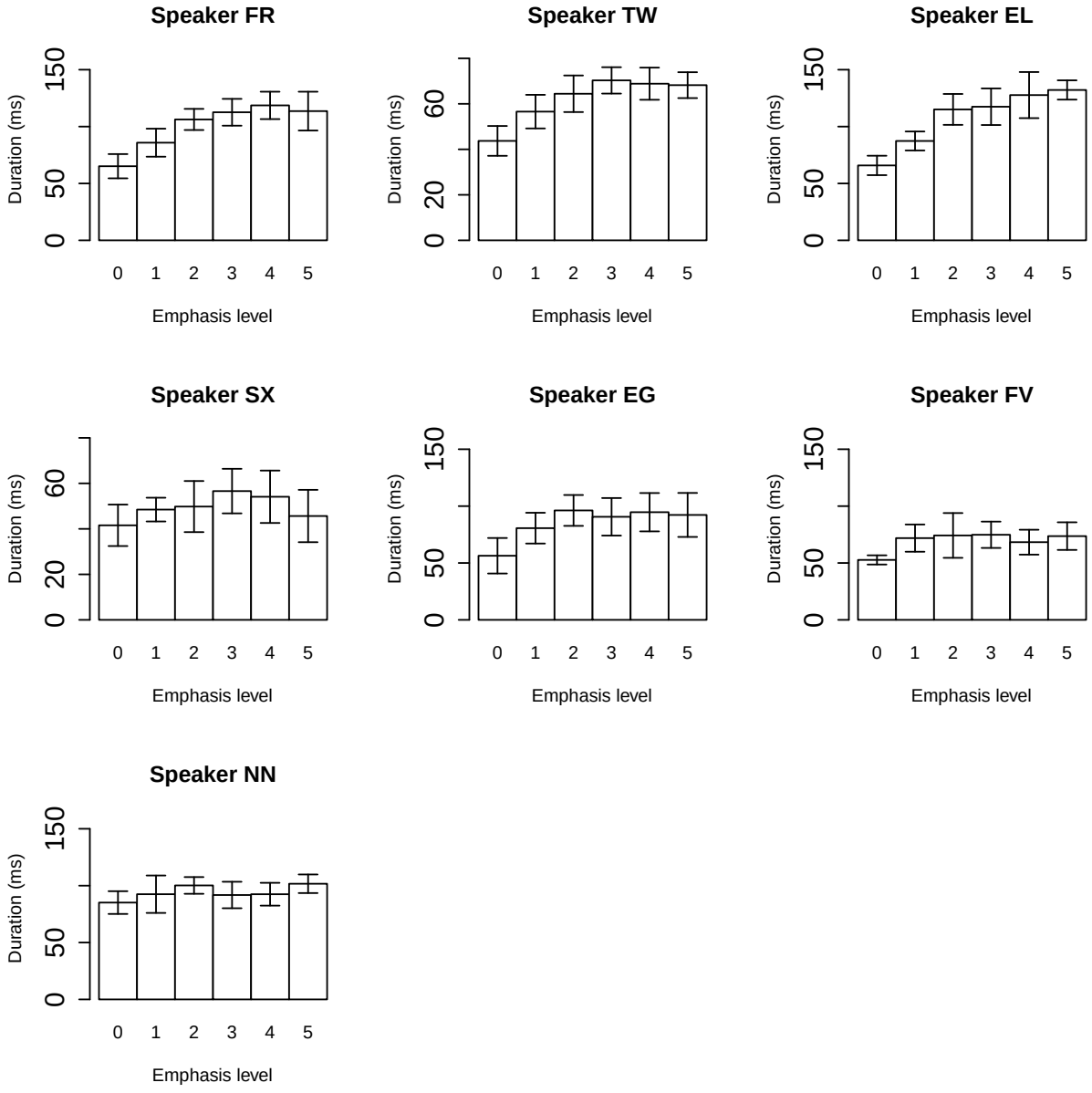

Figure 10: The average durations of preceding vowels for each emphasis level with 95% confidence intervals. The figures have different y-axis scales, because different speakers use different duration ranges.

In addition, we observe some differences in pre-consonantal vowel duration among different emphatic levels for some speakers, but the tendency is much less clear compared to the differences we observed in consonantal duration. Even the two speakers who showed the clearest six way differences in consonantal durations (Speakers FR and TW) do not show differences in preceding vowel duration among levels higher than 2. Speaker EL also does not show differences for levels higher than emphasis level 2. The next three speakers (Speaker SX, EG, FV) only show a two-way difference between plain consonants on the one hand and the emphatically lengthened consonants on the other, reflecting the general pattern found in Japanese (see above), but they do not show clear differences among the emphatically lengthened consonants. For Speaker NN, the difference between the plain and emphatically-lengthened conditions does not seem substantial.[9]

To compare the patterns of consonantal durations and vowel durations, Table 11 summarizes the correlation (Pearon's r) values between durations of preceding vowels and emphasis levels for each speaker. Recall that these values are a measure of the strength of correlation between durations and emphasis levels. In the rightmost column, r values for consonants are reproduced for each speaker from Table 10 for the sake of comparison.

Table 11: Summary of the correlation (Pearson's r) between duration of preceding vowels and emphasis levels for each speaker. The r values for consonants are reproduced in the rightmost column for the sake of comparison.

	Preceding Vowels		**Consonants**	
Speaker	***r***	**sig.**	***r***	**sig.**
Speaker FR	.52	$p < .001$	.91	$p < .001$
Speaker TW	.43	$p < .01$	.74	$p < .001$
Speaker EL	.65	$p < .001$	.70	$p < .001$
Speaker SX	-.007	*n.s*	.66	$p < .001$
Speaker EG	.16	*n.s*	.65	$p < .001$
Speaker FV	-.02	*n.s*	.56	$p < .001$
Speaker NN	.14	*n.s*	.21	$p < .001$

Table 11 confirms the observations we made regarding Figure 10. Speakers FR, TW and EL

[9] A post-hoc t-test shows that there are indeed no significant differences ($t(10) = .98, n.s.$).

all show positive correlation, but these *r*-values are not as high as those we observed for consonantal durations with these speakers (see the rightmost column). The other speakers do not show a significant correlation between emphasis levels and preceding vowel duration. It can be concluded that the consonant gemination pattern in Japanese targets consonant duration more than preceding vowel duration, and that the effect of lengthening is primarily localized to consonant duration.

4 General discussion

4.1 Theoretical discussion

There are a number of theoretical questions that the current results bear on. One implication of this study, beyond providing the first experimental description of the emphatic gemination pattern in Japanese, is that at least two speakers showed clear six way differences, and all but one speaker showed a steady linear correlation between duration and emphasis level. In general, then, articulatorily speaking, it is likely that speakers can make durational differences that go beyond a binary distinction. This finding is interesting in light of the observation that for lexical singleton-geminate contrasts, the differences are usually limited to a binary difference (Ladefoged and Maddieson, 1996). This conclusion—that they can produce more than binary durational differences—does not of course automatically generalize to the speakers of other languages, but at least some speakers from one language are able to produce six-way durational differences. Future cross-linguistic examinations are hoped for in order to further examine this conclusion.

Another theoretical question that arises is why, given that speakers can in general make durational distinctions beyond a binary contrast, do natural languages usually exploit only a two-way distinction for lexical contrasts. An obvious hypothesis would be that perception is playing a role here—a three way durational contrast may be difficult to perceive and may cause confusion, which is to be avoided, following the spirit of Adaptive Dispersion Theory (see e.g. Diehl et al. 2004; Flemming 1995; Liljencrants and Lindblom 1972; Lindblom 1986; Padgett 2002; Schwartz et al. 1997a,b; see especially Engstrand and Krull 1994 and Kawahara 2012 for the relevance of percep-

tual dispersion in durational dimensions). Another hypothesis is more formal—that phonological systems build on binary distinctive features (the length contrast being one of these) (Chomsky and Halle, 1968), such that lexical contrasts are always limited to binary distinctions by Universal Grammar. Settling this debate is beyond the scope of the present paper, and warrants future experimental studies.

Third, an anonymous reviewer raises the question of how the current results bear on the representation of long segments generally. In current phonological theory, there are, broadly speaking, two approaches: (i) simply representing long segments as [+long] (Chomsky and Halle, 1968; Kuroda, 1967), or (ii) separating timing slots from segmental content, with timing slots being represented by, for example, C-slots (McCarthy, 1979), moras (Hayes, 1989) or by root nodes (Selkirk, 1991) (see Davids 2011 for an overview). For the emphatic lengthening phenomenon at least, it is impossible to simply deploy a [+long] feature, because the contrast is not binary. The second type of theory can more parsimoniously accommodate the lengthening phenomenon, because in principle, timing slots can be added for each emphasis level.[10]

Finally, recall that for all the speakers, the emphatically lengthened consonants were longer than the plain consonants (despite the fact that not all speakers realized differences among the different levels of emphasis). Moreover, as observed in all the results figures, all speakers have a very large increase in duration from plain consonants to emphatically lengthened consonants, and this increase is larger than the observed differences among the different levels within the emphatically lengthened consonants.[11] It then seems that Japanese speakers overall make a binary contrast between plain and emphatically lengthened consonants, and within the emphatic consonants, speakers choose different options about how to scale the degrees of emphasis.[12] That is,

[10]We should perhaps be careful about extending this conclusion to lexical geminates, however, because there are no guarantees that lexical geminates and emphatically lengthened geminate are representationally identical. We note, however, that both lexical geminates and emphatically lengthened geminates in Japanese show lengthening of the preceding vowels. This parallel suggests that lexical geminates and emphatically lengthened geminates may have something in common with respect to their phonological representation.

[11]This observation is even more surprising when we consider the fact that Japanese is a mora-timed language in which the duration of each mora is more or less consistent (see Beckman 1982; Han 1994; Port et al. 1987; Warner and Arai 2001 for discussion). Since single level gemination is counted as the addition of one mora, at least in terms of orthography, the larger difference between plain consonants and consonants with level 1 emphasis is unexpected from the moraic point of view.

[12]An anonymous reviewer raises another interesting explanation, which is that singleton (non-emphatic) and gem-

plain consonants are not 0 on the scale of emphasis, but instead are categorically different from all emphatics. This pattern may imply that the distinction between non-emphatic and emphatic is semantically more important than the finer degrees of emphasis, and that speakers reflect this difference of semantic importance in their production of emphasized and plain consonants.

4.2 Remaining questions

The current study also raises several questions. For example, would speakers of other languages be able to make similar durational differences? Would there be a difference between languages that exploit duration-based lexical contrasts (as in Japanese) and those that do not (as in English)? In a follow-up study currently in progress, we have examined a similar phenomenon in English, using examples like *That guy is soooooo creepy* and *That joke is suuuuper funny* (Dresher et al., 2013). The preliminary results, partly illustrated in Figure 11, show that at least some English speakers—who do not have a phonological length contrast in their native language—do make differences that are similar to those found in our current project. This sort of cross-linguistic comparison should be elaborated in future studies.

The phenomenon of emphatic lengthening in Japanese (and other languages) can be studied from further perspectives as well. For example, it is conceivable that emphasis is conveyed along other acoustic dimensions, such as intensity differences. In such cases, would we expect to see the same sort of fine-grained patterns which go beyond binary distinctions? Another question is whether Japanese speakers make similar levels of differences in actual production patterns in naturalistic settings, i.e., even when they are not prompted in experimental settings. Yet another question is that of perception: given that speakers can produce distinctions that go beyond binary, to what extent can native speakers perceive them?[13] These are interesting questions, which are

inate (the first order emphatic) durations are already encoded in the Japanese phonology. Since the phonological contrast in Japanese is not three-way, the distinction between the singleton and geminate can be maximally realized within the acoustic space of (phonemic) stop length distinction, even in the emphasis context. Our experiment on a similar phenomenon in English shows (Dresher et al., 2013), however, that English speakers show the same pattern as the Japanese speakers (see Figure 11), indicating that even in a language that does not have a phonemic contrast, the difference between non-emphatic and emphatic is more robust than differences between different levels of emphasis.

[13] A perception experiment using English listeners is in progress.

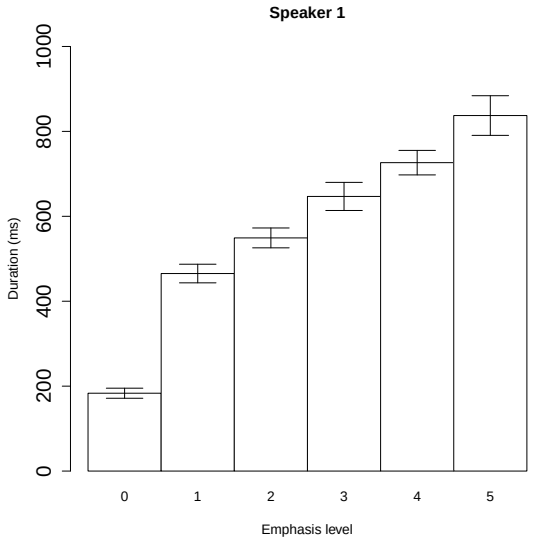

Figure 11: An illustration of vowel lengthening by emphasis level by an English speaker. Taken from Dresher et al. 2013.

however beyond the scope of the current study. In general, our project may raise more questions than it answers, but for that very reason, this project will open up opportunities for future studies on the phonetics of pragmatic emphatic lengthening.

4.3 Final summary

In summary, while the phonetics of lexical singleton-geminate contrasts is well studied, there are few if any phonetic studies on pragmatically lengthened segments. The current study has offered a first study of pragmatically lengthened consonants, using Japanese as a case study. It has shown that two speakers made six-way durational distinctions, and all other speakers (expect for one) showed a significant correlation between consonant duration and emphasis level. The locus of durational differences is chiefly localized to consonant duration. This study raises many questions about the phonetics of pragmatic lengthening in Japanese and other languages, and thus opens up opportunities for future phonetic studies.

References

Abel, Sharon. 1972. Discrimination of temporal gaps. *Journal of the Acoustical Society of America* 52: 519–524.

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

Aizawa, Yoshiko. 1985. Intensification by so-called "choked sounds"–long consonants–in Japanese. In *The study of sounds 21*, ed. The Phonetic Society of Japan, 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.

Bals Baal, Berit Anne, David Odden, and Curt Rice. 2012. An analysis of North Saami gradation. *Phonology* 29 (2): 165–212.

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

Bellem, Alex. 2007. Towards a comparative typology of emphatics: Across Semitic and into Arabic dialect phonology. Doctoral dissertation, SOAS.

Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5 (9/10): 341–345.

Boersma, Paul, and David Weenink. 1999–2013. Praat: Doing phonetics by computer. Software.

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

Cedrus Corporation. 2010. Superlab v. 4.0. Software.

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

Cohn, Abigail, William Ham, and Robert Podesva. 1999. The phonetic realization of singleton-geminate contrasts in three languages of Indonesia. *Proceedings of ICPhS* XIV: 587–590.

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

DiCanio, Christian T. 2012. The phonetics of fortis and lenis consonants in Itunyoso Trique. *International Journal of American Linguistics* 78 (2): 239–272.

Diehl, Randy, Björn Lindblom, and Carl Creeger. 2004. Increasing realism of auditory representations yields further insights into vowel phonetics. *Proceedings of ICPhS* XV: 1381–1384.

Dresher, Natalie, Sarah Korostoff, Melanie Pangilinan, Jessica Trombetta, and Shigeto Kawahara. 2013. The phonetics of emphatic vowel lengthening in English. Talk presented at Rutgers Uni-

versity Aresty Symposium.

Engstrand, Olle. 1987. Preaspiration and the voicing contrast in Lule Sami. *Phonetica* 44: 103–116.

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.

Flemming, Edward. 1995. Auditory representations in phonology. Doctoral dissertation, University of California Los Angeles.

Fujisaki, H., K. Nakamura, and T. Imoto. 1975. Auditory perception of duration of speech and non-speech stimuli. In *Auditory analysis and perception of speech*, eds. G. Fant and M. A. A. Tatham, 197–219. London: Academic.

Fukui, Seiji. 1978. Nihongo heisaon-no encho/tanshuku-niyoru sokuon/hisokuon toshite-no chooshu [Perception for the Japanese stop consonants with reduced and extended durations]. *Onsei Gakkai Kaihou [Phonetic Society Reports]* 159: 9–12.

Games, Sara. 1976. *Quantity in Icelandic: Production and perception*. Hamburg: Helmut Buske Verlag.

Gordon, Matthew, Pamela Munro, and Peter Ladefoged. 2000. Some phonetic structures of Chickasaw. *Anthropological Linguistics* 42: 366–400.

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

Han, Mieko. 1962. The feature of duration in Japanese. *Onsei no Kenkyuu [Studies in Phonetics]* 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.

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

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

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

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

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.

Hirato, Nobuo, and Shinichiro Watanabe. 1987. Sokuon-no chikaku-to kouzoku boin-no jizoku jikan tono kankei [The relationship between the perception of geminates 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. *Proceedings of ICPhS* XVI: 909–912.

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

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.

Jaeger, Jeri J. 1983. The fortis/lenis question: Evidence from Zapotec and Jawoñ. *Journal of Phonetics* 11: 177–199.

Kahn, M. 1975. Arabic emphatics: The evidence for cultural determinants of phonetic sex-typing. *Phonetica* 31: 38–50.

Kato, Hiroaki, and Minoru Tsuzaki. 1994. Intensity effect on discrimination of auditory duration flanked by preceding and suceedings tones. *Journal of the Acoustical Society of Japan* 15 (5): 349–351.

Kato, Hiroaki, Minoru Tsuzaki, and Yoshinori Sagisaka. 1997. Acceptability for temporal modification of consecutive segments in isolated words. *Journal of the Acoustical Society of America* 101: 2311–2322.

Kato, Hiroaki, Minoru Tsuzaki, and Yoshinori Sagisaka. 2002. Effects of phoneme class and duration on the acceptability of temporal modifications in speech. *Journal of the Acoustical Society of America* 111: 387–400.

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

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

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. 2013. Emphatic gemination in Japanese mimetic words: A wug-test with auditory stimuli. *Language Sciences* 40: 24–35.

Kawahara, Shigeto. to appear. The phonetics of *sokuon*, obstruent geminates. In *The handbook of*

Japanese language and linguistics: Phonetics and phonology, ed. Haruo Kubozono. Mouton.

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

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.

Krähenmann, Astrid. 2003. *Quantity and prosodic asymmetries in Alemannic*. Berlin: Mouton de Gruyter.

Kuroda, S. Y. 1967. *Yawelmani phonology*. Cambridge, MA: MIT Press.

Ladefoged, Peter, and Ian Maddieson. 1996. *The sounds of the world's languages: 2nd edition*. Oxford: Blackwell Publishers.

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

Laufer, Asher, and Thomas Baer. 1988. The emphatic and pharyngeal sounds in Hebrew and in Arabic. *Language and Speech* 31 (2): 181–205.

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

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 *Experimental phonology*, eds. John Ohala and Jeri Jaeger, 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.

Local, John, and Adrian Simpson. 1999. Phonetic implementation of geminates in Malayalam nouns. *Proceedings of ICPhS* XIV: 595–598.

Maddieson, Ian. 1997. Phonetic universals. In *The handbook of phonetics sciences*, eds. W. Hardcastle and J. Laver, 619–639. Oxford: Blackwell.

McCarthy, John J. 1979. Formal problems in Semitic phonology and morphology. Doctoral dissertation, MIT. Published by Garland Press, New York, 1985.

McCarthy, John J. 1994. The phonetics and phonology of Semitic pharyngeals. In *Papers in laboratory phonology III: Phonological structure and phonetic form*, ed. Patricia A. Keating, 191–233. Cambridge: Cambridge University Press.

McKay, Graham. 1980. Medial stop gemination in Rembarrnga: A spectrographic study. *Journal of Phonetics* 8: 343–352.

Nasu, Akio. 1999. Chouhukukei onomatope no kyouchou keitai to yuuhyousei [Emphatic forms of reduplicative mimetics and markedness]. *Nihongo/Nihon Bunka Kenkyuu [Japan/Japanese*

Culture] 9: 13–25.

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

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

Ofuka, Etsuko, Yoko Mori, and Shigeru Kiritani. 2005. Sokuon-no chikaku-ni taisuru senkoo-kouzoku boincho-no eikyoo [The effects of the duration of preceding and following vowels on the perception of geminates]. *Onsei Kenkyu [Journal of the Phonetic Society of Japan]* 9: 59–65.

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

Padgett, Jaye. 2002. Contrast and postvelar fronting in Russian. *Natural Language and Linguistic Theory* 21: 39–87.

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

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

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.

Prince, Alan. 1980. A metrical theory for Estonian quantity. *Linguistic Inquiry* 11: 511–562.

R Development Core Team. 1993–2013. *R: A language and environment for statistical computing*. Vienna, Austria. R Foundation for Statistical Computing. Software, available at http://www.R-project.org.

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

Schwartz, Jean-Luc, Louis-Jean Boë, Nathalie Valleé, and Christian Abry. 1997a. The dispersion focalization theory of vowel systems. *Journal of Phonetics* 25: 255–286.

Schwartz, Jean-Luc, Louis-Jean Boë, Nathalie Valleé, and Christian Abry. 1997b. Major trends in vowel system inventories. *Journal of Phonetics* 25: 233–253.

Selkirk, Elisabeth. 1991. A two-root theory of length. In *University of Massachusetts occasional papers in linguistics 14: Papers in phonology*, eds. E. Dunlap and Jaye Padgett. Amherst: GLSA.

Shrotriya, N., A. S. Siva Sarma, R. Verma, and S. S. Agrawal. 1995. Acoustic and perceptual characteristics of geminate Hindi stop consonants. *Proceedings of ICPhS* XIII: 132–135.

Takeyasu, Hajime. 2012. Sokuon-no chikaku-ni taisuru senkoo onsetsu shiin boin-no jizokujikan-no eikyoo [Effects of the consonant and vowel durations in the preceding syllable on the perception of geminate stops in Japanese]. *On'in Kenkyu [Phonological Studies]* 15: 67–78.

Tserdanelis, Georgios, and Amalia Arvaniti. 2001. The acoustic characteristics of geminate consonants in Cypriot Greek. *Proceedings of the Fourth International Conference on Greek Linguistics: 29-36*.

Warner, Natasha, and Takayuki Arai. 2001. Japanese mora-timing: A review. *Phonetica* 58: 1–25.