音声研究　第19巻第2号
2015(平成27)年8月
33–43頁

Journal of the Phonetic Society
of Japan, Vol. 19 No. 2
August 2015, pp. 33–43

特集論文

Edge Prominence and Declination in Japanese Jaw Displacement Patterns: A View from the C/D Model

Shigeto Kawahara*, Donna Erickson** and Atsuo Suemitsu***

日本語における句内部の顎の動き：C/Dモデルの観点から

SUMMARY: The dominant view in the field of Japanese phonetics and phonology is that Japanese metrical prominence, if anything, manifests itself as pitch accent, whose primary acoustic correlate is F0 fall. Work by Osamu Fujimura has challenged this view, by arguing that Japanese has stress, which is realized by way of increases in jaw opening. In addition, he argues that jaw displacement patterns show declination within a phrase, just as F0 does. This paper reports an experiment using EMA (ElectroMagnetic Articulograph) which examined these claims. The results of the current experiment show that these claims by Fujimura are in principle correct empirically, and hold across all six native speakers of Japanese tested in this experiment. In addition, the current results reveal that Japanese exhibits final stress, which is a new finding going beyond the original insights offered by Fujimura's work. A further acoustic analysis shows that initial and final stress manifest itself in high F1, and surprisingly, low intensity. All in all, we conclude that Japanese has both initial and final stress, with declination observed within the phrase-internal syllables.

Key words: Japanese, stress, jaw, Fl, intensity, phrasing

1. Introduction

Human speech streams are not merely a monotonic sequence of segments or syllables; human utterances are instead grouped into smaller rhythmic units or phrases, and within each phrase, some elements receive prominence, whereas others do not. We refer to such prominence as "metrical prominence." In Metrical Phonology (Liberman and Prince 1977 *et seq.*), syllables within a foot are assigned the status of either "strong" or "weak." The dominant view in the field of Japanese phonetics and phonology has been that Japanese metrical prominence, if anything, realizes itself in terms of lexical pitch accent, where the primary acoustic correlate of pitch accent is F0 movement[1]. We thus observe a typical statement of the following sort: "Japanese is a pitch accent language, whereas English is a stress-based language" (e.g. Kawahara 2015, McCawley 1978). What is implied in this statement is that Japanese does not have stress[2].

Osamu Fujimura's C/D model has long been challenging this view (Fujimura 1999, 2000, 2001, 2003), although, in our opinion, this claim has been under-appreciated[3]. In particular, he claims that (i) Japanese has stress and that (ii) stress manifests itself in the amount of the jaw displacement, in such a way that stressed syllables show larger jaw displacement (Fujimura 2001, p. 172, Fujimura 2003, p. 227). Let us consider a sample syllable triangle representation, shown in Figure 1, adapted from Fujimura (2003, p. 228). This diagram is for three Japanese sentences with three different syntactic subjects with different pitch accent (/hasi-ga aru/ (unaccented) 'An edge exists'; /hasi'-ga aru/ 'A bridge exists'; /ha'si-ga aru/ 'A chopstick exists') (in this paper, following Fujimura and other work, we use phonemic transcription rather than narrow phonetic transcriptions to represent Japanese sounds). In the C/D model, all of these sentences are mapped onto the same sequence of syllable triangles, shown at the bottom of the figure.

Now recall that in the C/D model, the height of syllable triangles represents metrical prominence, one of

* The Keio Institute of Cultural and Linguistic Studies（慶應義塾大学言語文化研究所）
** Kanazawa Medical University（金沢医科大学）
*** Japan Advanced Institute of Science and Technology（北陸先端科学技術大学院大学）

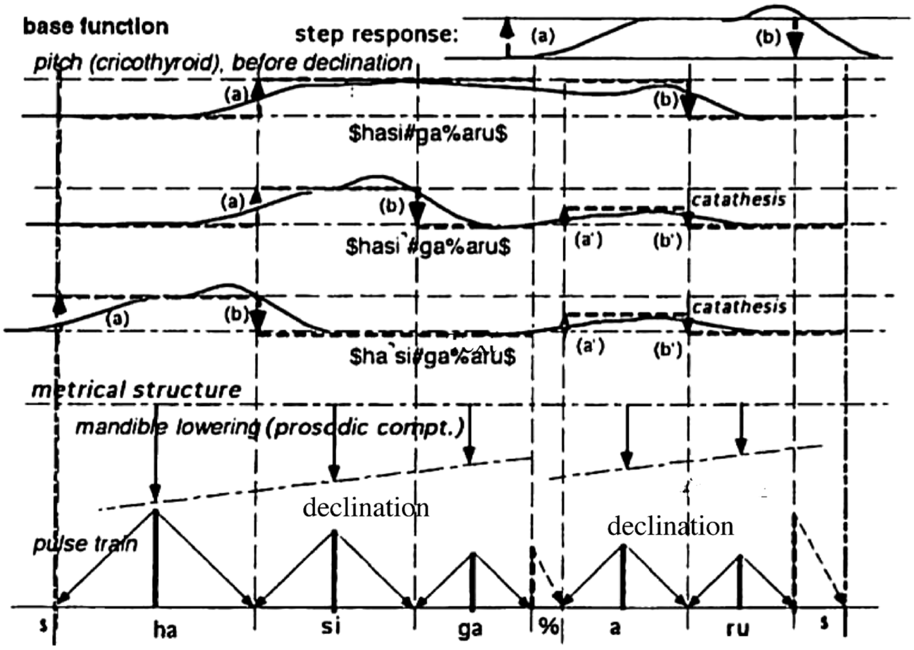

Figure 1 The syllable triangle representation of three sentences with different syntactic subject nouns with three different pitch accent patterns (/hasi/ 'edge,' /hasi'/ 'bridge,' /ha'si/ 'chopstick'). The top panel shows the idealized F0 patterns of the three sentences. The bottom panel shows the syllable triangle representations, whose height corresponds to jaw displacement magnitudes. This figure is adapted from Fujimura (2003, Figure 2, p. 228). Spelling errors in the original text are corrected.

whose phonetic correlate is jaw displacement (Erickson 2002, Fujimura 2000, 2002, 2007). Thus Figure 1 makes three specific predictions, which are shown in (1).

(1) Three theses about Japanese phrasing and stress
  a. The three sentences are all grouped into two phrases: (hasiga) and (aru).
  b. Within each phrase, the initial syllable is the strongest.
  c. The height of triangles generally declines within a phrase, similar to F0 declination.

Perhaps (1)a is not so controversial; /hasiga/ is the syntactic subject of the sentence, and the /aru/ is the syntactic predicate (or a verb phrase). (1)b is probably more likely to come as a surprise, because it implies that Japanese has phrase-initial stress. (1)c is also new and interesting; although it is uncontroversial that Japanese F0 declines over the course of an utterance (Fujisaki and Hirose 1984, Kawahara and Shinya 2008, Pierrehumbert and Beckman 1988, Poser 1984), the C/D model claims that there is a similar sort of declination in jaw displacement patterns[4)].

The goal of our current research is to test these predictions, or assertions, of the C/D model. Although the questions are framed from the perspectives of the C/D model, the questions addressed in this study can be stated in more theory-neutral terms, as in (2). Therefore, although the current research is much inspired by the C/D model, the implications of this study are not limited to the domain of that particular theory.

(2) The questions stated more generally
  a. Does Japanese metrical prominence manifest itself in some dimension other than lexical pitch accent?
  b. Does jaw movement show initial prominence within a phrase?
  c. Does jaw movement show declination within a phrase?

2. Method

To address the assertions summarized in (1) and (2), the current experiment deployed a 3D EMA (Electro-Magnetic Articulograph), which allows us to quantitatively measure the amount of jaw displacement. The current experiment used Carstens AG500, and it took place in a sound-attenuated laboratory at Japan Advanced Institute of Science and Technology (JAIST). Parts of the results that follow were reported previously in Kawahara et al. (2014a), but the current paper substantially augments Kawahara et al. (2014a) with data of three additional speakers, and reinterprets the results of Kawahara et al. (2014a) from the perspective of the C/D model. This paper also reports acoustic analysis of the data, which Kawahara et al. (2014a) did not do.

2.1 Stimuli

Since we know that vowel quality independently affects degrees of jaw displacement to substantial degrees (Kawahara et al. 2014b, Keating et al. 1994, Recasens 2012, Williams et al. 2013), we controlled the vowels in the stimulus sentences. In particular, all the syllables in the stimuli consisted of /a/. In this paper, we focus on two stimulus sentences: /akagasa-da/ 'That's a red umbrella' and /akapajama-da/ 'That's a red pajama' (/j/ here represents a voiced palatal fricative, not a palatal glide). Kawahara et al. (2014a) report longer sentences produced by three speakers, but the current paper focuses on these two sentences due to space limitation.

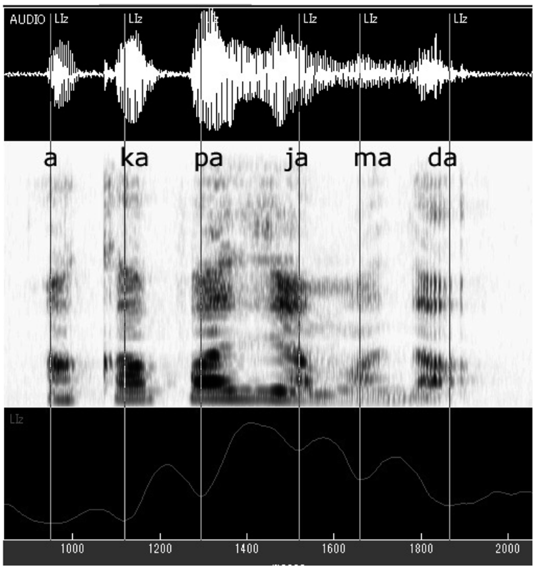

Figure 2 A sample jaw movement data, illustrating our measurement procedure (mview screen). The bottom panel shows the jaw displacement pattern. Vertical lines show our measurement points. This figure is based on a token of /akapajamada/ produced by Speaker T.

2.2 Speakers

Six native speakers of Japanese participated in this experiment: Speaker K (female), Speaker H (female), Speaker S (male), Speaker N (female), Speaker T (female), and Speaker Y (female). Speaker S is the first author of the paper, whose data we will interpret with caution. The behaviors of Speakers K, H, S were previously reported in Kawahara et al. (2014a). No speakers except for Speakers S and N were aware of the purpose of the experiment. All speakers were in their 20's or 30's at the time of recording.

2.3 Procedure

For EMA measurement, one sensor was placed on the lower medial incisors in order to track the jaw lowering movement. Four additional sensors were used to establish references in order to account for the effects of head movement. A biteplate with three additional sensors was used to estimate the location of the occlusal plane. The distance between the biteplate and the sensor on the lower incisors was calculated as a measure of the degree of jaw displacement. The articulatory and acoustic data were digitized at sampling rates of 200 Hz and 16 kHz, respectively. The stimuli, together with stimuli for other experiments, were presented in a randomized order on a Powerpoint screen, which was presented in front of the speaker. The speaker read each stimulus 6 times, but there were sometimes tracking errors or mispronunciation errors.

2.4 Measurement: Articulation

The analysis of jaw movement was conducted using custom software (mview), developed by Haskins Laboratories. A sample analysis screen is shown in Figure 2. The top panel shows the waveform, the middle panel shows the spectrogram, and the contour in the bottom panel shows the jaw movement. Within each vowel, the amount of jaw displacement was measured at the point of maximum jaw displacement, shown with vertical lines in Figure 2.

Occasionally, some speakers—especially Speaker H and Speaker Y—showed maximum jaw opening well before the start of the vowel of the first syllable of the utterance. We limited our measurement to jaw lowering movements that occur during the intervals when the vowels are clearly present in the acoustic signal. This choice was to be conservative: we wanted to make sure that our measurements are those that are related to linguistic signals with clear acoustic energy; also, the

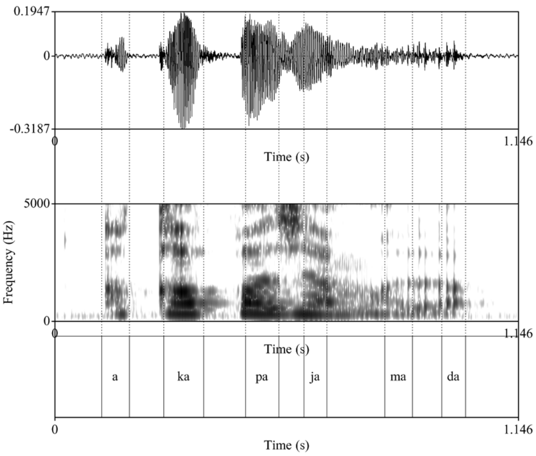

Figure 3 An illustrative figure which shows the acoustic measurement protocol. Vocalic intervals are annotated with their whole syllable content. This figure is based on the token of /akapaja-mada/ produced by Speaker H.

magnitudes of initial vowels would be smaller in our method, allowing us to examine the prediction of the C/D model in a more conservative way.

All speakers read both of the stimulus sentences 6 times, except for Speaker H who read the "akapa-jamada" sentence 7 times. No extreme outliers were found. Error bars are shown in the following illustrative figures in the result section to show the degree of variability found in the current data.

2.5 Measurement: Acoustics

In order to also investigate the possible acoustic effects of metrical prominence in Japanese, acoustic analysis was conducted. In particular, we would expect F1 to be higher when the jaw is lower, because the Helmholtz resonation—the source of F1—gets higher as the oral cavity is more open (Erickson et al. 2012, Johnson 2003, Stevens 1998). We might also expect stressed vowels to be more intense, because large jaw opening would emit more energy out of the oral cavity (Schulman 1989). This prediction is expected from a cross-linguistic consideration as well; e.g. English stressed vowels show higher intensity (Plag et al. 2011). F0 was not measured, because the vowels in the current stimuli differ in their accentual properties, and hence would not be very informative.

For each token, vocalic boundaries are placed based on the cessation of formant structures, as illustrated by the sample spectrogram shown in Figure 3. A Praat script (Boersma 2001, Boersma and Weenink 1999–2015) was written to automatically detect the midpoint of each vowel interval, and create a 30 ms window centering around that midpoint. The average values of F1 and intensity were automatically extracted within each 30 ms window. The figures in the results below were all produced using R (R Development Core Team, 1993–2015).

3. Results: Articulatory Data

3.1 Individual Patterns

Let us first consider the jaw displacement pattern of Speaker K, shown in Figure 4. Here and throughout, we use a bar graph instead of a line graph, because a bar graph representation is more in line with the general spirit of the C/D model in that a bar directly corresponds to the height of a syllable triangle. The error bars are standard errors.

The left panel of Figure 4, which shows her jaw displacement pattern of /akagasada/, exhibits a very clear strong jaw opening initially, and the amount of jaw opening generally declines until the penultimate syllable /sa/. These patterns are exactly as those expected from the C/D model perspective, as reviewed in the introduction. One new finding, which was also alluded to in Erickson et al. (2014) and Kawahara et al. (2014a), was that the final syllable shows large jaw opening as well. We take this as evidence that Japanese has final stress, in addition to initial stress[5)].

Looking at the right panel of Figure 4, which shows the jaw movement of /akapajamada/, we observe that the patterning is generally the same as that of /akagasada/, except that the initial two syllables show prominence. In a sense, it looks as if the strength of the initial syllable "spills over" to the second syllable. We can consider this as resulting from the coarticulation of the second syllable to the initial syllable. Another idea we can entertain, although we are not ready to defend it in full detail, is that the first two syllables constitute a foot in Japanese, and the whole foot receives some metrical prominence. At any rate, the difference between the left panel and the right panel in Figure 4 is something to be explained, which we are unfortunately unable to do in this paper.

Let us move on to the pattern of Speaker S, who is the first author of this paper. As shown in Figure 5, his patterns show three characteristics that we observed for Speaker K: (i) initial syllables show the large jaw opening, (ii) degrees of jaw opening decline until the penultimate syllables, and (iii) the final syllables show

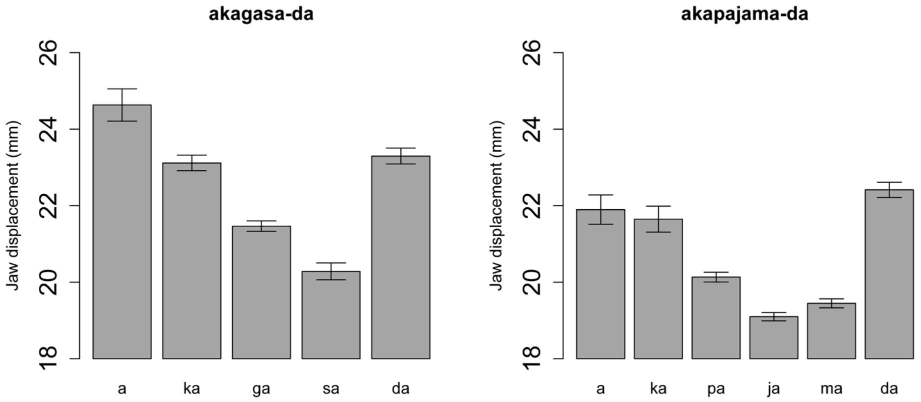

Figure 4 Jaw displacement pattern of Speaker K.

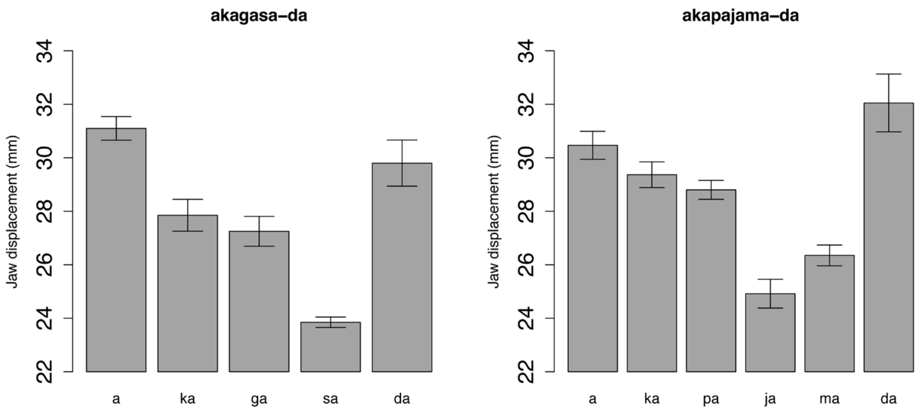

Figure 5 Jaw displacement pattern of Speaker S.

prominence.

An additional notable aspect is the observation that in the right panel, the antepenultimate syllable /ja/ is lower than the penultimate syllable /ma/. If declination alone dictates the jaw lowering pattern, /ma/ should be lower than /ja/. This observed reversal may have arisen from the fact that obstruent onsets may inhibit more jaw opening than sonorant onsets (Kawahara et al. 2014b, Vatikiotis-Bateson and Kelso 1993)—recall that /j/ is a fricative. We admit that this issue needs to be addressed in further detail in future studies[6].

One may quibble that the patterns in Figure 5 may have resulted from the bias that Speaker S had about this project. While we fully acknowledge that this is a legitimate concern, it is important to note that the other speakers show similar jaw displacement patterns. Moving on, for example, Speaker H, who was naive to the purpose of the project, shows almost identical jaw displacement patterns to those observed in Figure 5, as in Figure 6. One difference is that Speaker H showed large jaw opening in the first two syllables in the /akapajamada/ sentence. However, even this feature was observed in the speech of Speaker K (see Figure 4).

Moving on, Speaker N (Figure 7) is an "ideal" speaker who instantiates the crucial observations in this experiment without any additional complication, (i) initial and final syllables show large jaw opening, and (ii) there is declination in medial syllables, notwithstanding the fact that she showed a "reversal" between /ja/ and /ma/. We admit, however, that this speaker is a professional phonetician, and was familiar with the previous work of ours (Erickson et al. 2014, Kawahara et al. 2014a, 2014b) as well as with the C/D model in general.

Speaker T, although not as clear as the previous four speakers, shows the patterns that are similar to the other speakers we have examined above (Figure 8): the initial syllables (or the first two syllables) show the largest

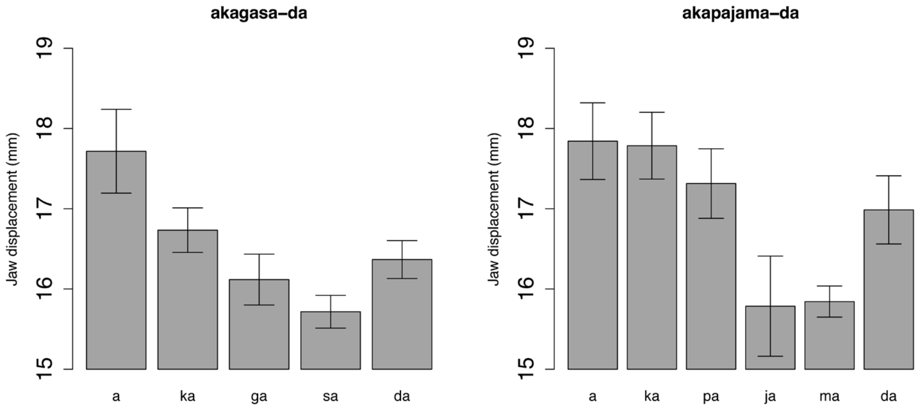

Figure 6　Jaw displacement pattern of Speaker H.

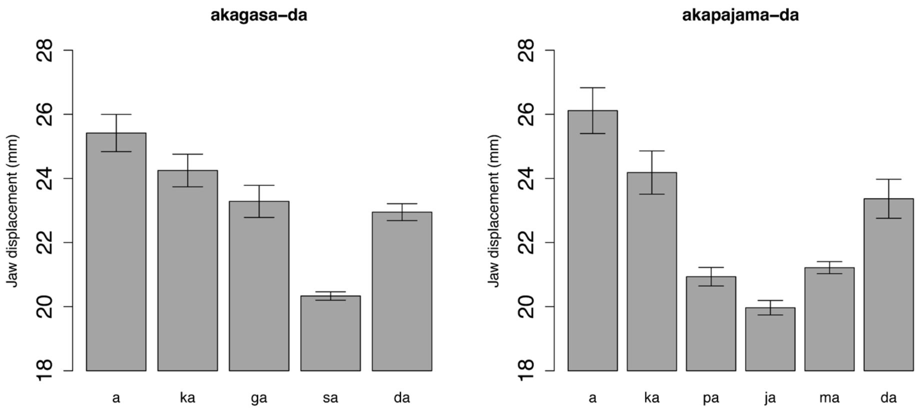

Figure 7　Jaw displacement pattern of Speaker N.

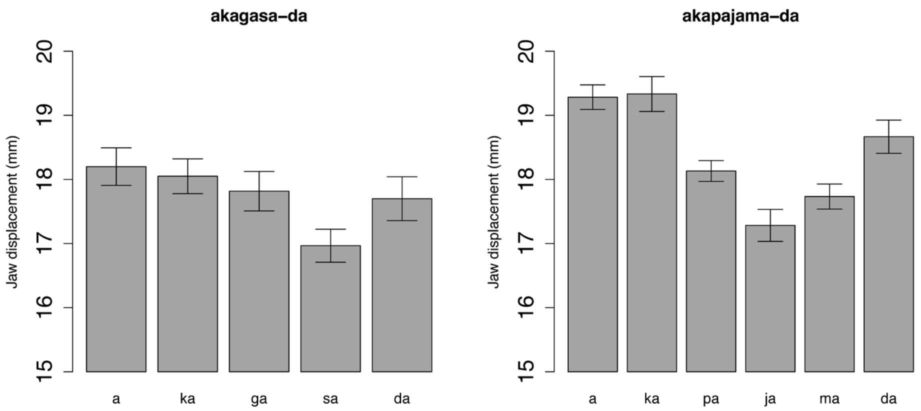

Figure 8　Jaw displacement pattern of Speaker T.

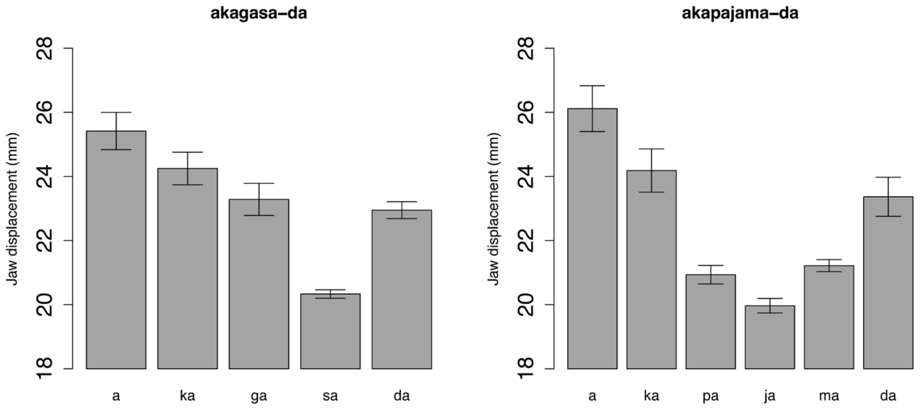

Figure 9 Jaw displacement pattern of Speaker Y.

jaw opening, after the declination, the final syllables show the strongest jaw opening.

The final speaker, Speaker Y, is as good as Speaker N in instantiating the general patterns of the current experiment, as shown in Figure 9: initial and final syllables show strong prominence, and there is a declination within phrase-medial syllables. As with some of the other speakers we observed above, there is a "reversal" between /ja/ and /ma/.

3.2 Discussion about the Articulation Patterns

As we have seen in this section, there are remarkably consistent jaw movement patterns across the speakers. Despite the fact that all the vowels are /a/, the jaw movement pattern is neither flat nor random. We reiterate that Speakers K, H, T, Y were unaware of the purpose of the current study. Given below is the summary of our findings.

(3) Summary of the articulatory patterns

a. Initial syllables show prominent jaw opening.

b. Jaw opening generally declines in subsequent syllables, although the second syllable may be as strong as the first syllable for some speakers.

c. The final syllables show large jaw opening, just like initial syllables.

d. (/ja/ can show smaller jaw opening than /ma/ despite the former appearing before the latter.)

The predictions of the C/D model are in line with what we actually observed in the current experiment, one sole exception is the final syllables, which consistently show large jaw opening. However, the general thesis—that Japanese has stress—remains valid. In fact, our results show that Japanese may have both initial and final stress.

There were some complications too, however. For example, as exemplified by the patterns of Speakers K and H, the second syllable can be almost as strong as the initial syllable. Contrary to the declination hypothesis, sometimes /ja/ shows smaller jaw opening than /ma/, despite the fact that the former appears before the latter.

4. Results: Acoustics

4.1 Individual Patterns

While the C/D model is primarily concerned with modeling the articulatory gestures, one would wonder what the acoustic correlates of phrasal prominence we observed in the previous section would be. To this end, this section reports some results of our acoustic analysis. Due to space limitation, we discuss only a subset of speakers, Speakers S, H and Y.

Figure 10 shows the acoustic properties—F1 on the top panel and intensity on the bottom panel—of Speaker S's speech. In both sentences (left and right figures), F1 is high in the sentence-initial and sentence-final position. This observation makes sense because jaw opening and F1 should be positively correlated with each other (Erickson et al. 2012, Johnson 2003, Stevens 1998). A bit surprising is the pattern of intensity: Initial and final vowels are less intense than phrase-medial vowels. This result is very counter-intuitive, because stressed vowels are usually louder than unstressed vowels in languages like English (Plag et al. 2011), and after all we are talking about metrical "prominence". Looking back at the sample waveform and spectrogram given in Figure 3, we do observe that the initial and final vowels are less intense.

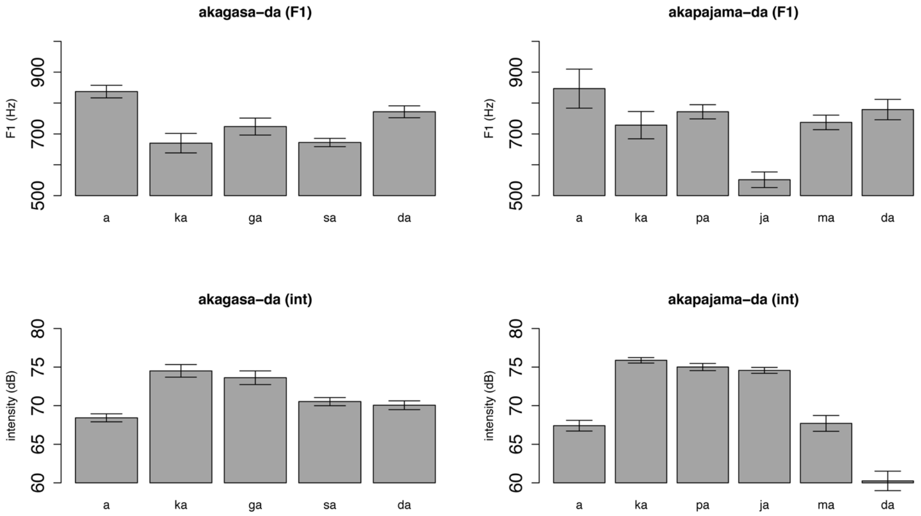

Figure 10　F1 and intensity patterns of Speaker S's speech.

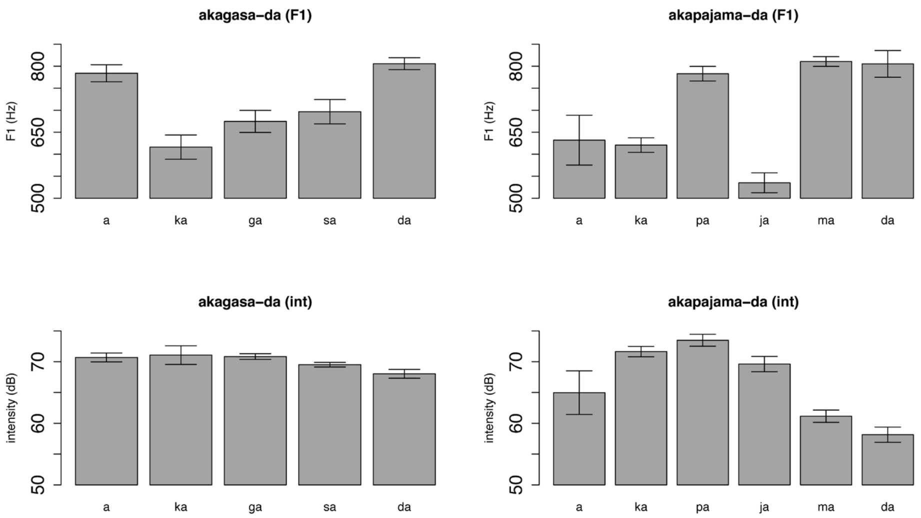

Figure 11　F1 and intensity of Speaker H's speech.

In Japanese, it is known that phrase-final vowels are creaky and weak (Kawahara and Shinya 2008), but the initial vowel's weakness cannot be attributed to creakiness[7]. However, we note that the fact that neither initial nor final vowels are acoustically prominent may be the reason why many previous studies have assumed that Japanese does not possess stress.

Speaker S's pattern is not anomalous: Speaker H shows several of the characteristics that Speaker S shows, as shown in Figure 11. Regarding F1, it is especially clear in the /akagasada/ sentence that the initial and the final vowels show higher F1. The tendency is less clear in the /akapajamada/ sentence, because the initial syllable does not show high F1, although we do observe high F1 for the last two syllables[8]. The intensity patterns are similar to that of Speaker S: initial and final vowels are weak, with intermediate vowels comparatively louder—the pattern is particularly clear in the /akapajamada/ sentence.

The patterns of Speaker Y, shown in Figure 12, confirm the previous observations: for F1, the initial and final vowels are generally high, even though, like

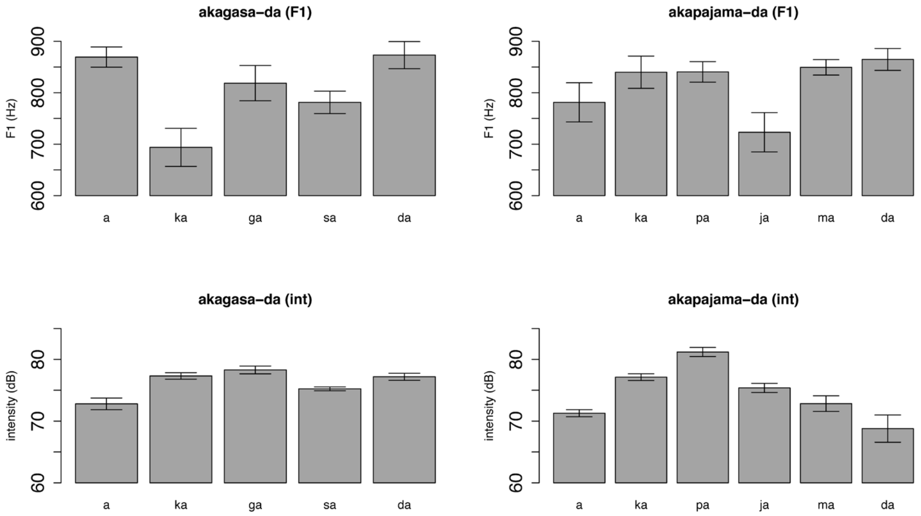

Figure 12 F1 and intensity of Speaker Y's speech.

Speaker H, the initial vowel of the /akapajamada/ sentence did not show high F1. The intensity contour shows that the initial and final vowels are usually weak, even though there is a reversal for the final syllable in the /akagasada/ sentence.

4.2 Discussion

All in all, "stressed" syllables in Japanese, which involve large jaw opening, are characterized by high F1 and low intensity. To what extent this stress is actually perceived by Japanese listeners is an interesting question. We do not mean to claim that listeners directly hear the jaw open more; however, it is possible—or even plausible—that F1 is a salient cue to stress, as is the case for English speakers (e.g. Mo et al. 2009). We do not have direct perceptual evidence yet that Japanese speakers use F1 to identify what we have called phrasal stress; however, we independently know that F1 is also a cue used by Japanese speakers to identify contrastive focus (e.g. Erickson 2004, Erickson et al. 2002, Maekawa 1997). Given these considerations, it is only logical to speculate that Japanese speakers use F1 to identify phrasal stress, although we reiterate that this claim needs to be tested experimentally in future studies. It may be that indeed Japanese stress is hard to hear, and that may be why Fujimura's claim that Japanese has stress has been under-appreciated. Nevertheless, when we examine jaw movement carefully, we do observe some sort of stress in the form of large jaw opening and increased F1.

5. Conclusion

The current EMA study has shown that Japanese has large jaw opening in initial and final syllables. In addition, within medial syllables, there is a general declination of jaw opening patterns. These observations are in line with the prediction of the C/D model, except that the model does not posit final stress in Japanese. Acoustically speaking, initial and final syllables are characterized with high F1 and non-intense vowels, although these patterns may not be as clear-cut as the articulatory patterns. From the perspectives of acoustics, Japanese stress is different from English stress. Nevertheless, in terms of jaw displacement at least, Japanese has stress phrase-initially and phrase-finally.

While the present data are in line with the predictions of the C/D model, there are a few questions that remain. First, why do Speakers H and T show a comparable degree of prominence in the first two syllables instead of having only the first one be most prominent? Is it due to coarticulation between the first two syllables or due to these syllables being parsed into a strong foot? Second, why do we sometimes observe a reversal between /ja/ and /ma/, contrary to what is expected from the declination hypothesis. Can we attribute this reversal to the independent observation that an obstruent onset inhibits jaw opening more than a sonorant onset? Third, why do initial syllables not show high intensity despite their large jaw opening? To answer this question, perhaps air pressure/air flow measurements are needed. Finally, we focused on jaw opening in this paper, because jaw movement is taken to be the primary manifestation of stress in the C/D model. However, how Japanese stress affects tongue movements, and how they affect the acoustics of Japanese stress are yet to be investigated. These are all interesting questions that need to be pursued in future research.

Acknowledgements

This work was supported by NSF IIS-1116076, NIH DC007124, and the Japan Society for the Promotion of Science, Grants-in-Aid for Scientific Research (A) #25240026 and (C) #25370444.

Notes

1) Indeed, many Japanese phonologists assume that accent is assigned to the head syllable of a metrical foot (e.g. Ito and Mester 2012, Kawahara 2015, Kubozono 2003, Shinohara 2000). This view is, we believe, not necessarily incompatible with the idea that Japanese also exhibits stress as its metrical prominence, since the control of F0 and jaw movement can be more or less independent. See Kawahara et al. (2014a) for articulatory data which shows that pitch accent does not substantially affect jaw displacement patterns.

2) In the previous studies of Japanese, even sentential focus, such as that observed in contrastive focus or in wh-elements, has been primarily studied in terms of f0 (in terms of prosodic phrasing or phonetic realization of lexical accent) (Ishihara 2003, Nagahara 1994, Pierrehumbert and Beckman 1988, Sugahara 2003). On the other hand, studies on English prominence associated with contrastive focus have revealed that jaw displacement is an important articulatory correlate of contrastive emphasis (see Erickson 2002, Erickson et al. 1998, Menezes et al. 2003 and references cited therein). Erickson et al. (2000) is a rare attempt to look at the effect of contrastive focus on jaw movement both in English and Japanese, but their results are admittedly preliminary. Examining the effect of contrastive focus on jaw movement in Japanese is an important topic for future research.

3) At the end of the paper, we conjecture why his claim may have been under-appreciated.

4) These three theses are reiterated in Figure 1.8 of Fujimura (2007, p. 18), a book on the C/D model written in Japanese. In that figure, declination of mandible lowering within each phrase is shown with a separate tier. In the main text, Fujimura (2007, pp. 19–20) states that "within each accentual phrase, syllable magnitude generally declines, and at the beginning of the next phrase, syllable magnitude gets large and declines again" (our translation).

5) Erickson et al. (2014) found this final prominence in the L2 English speech by Japanese speakers as well. See also Kawahara et al. (2014a) for evidence that this final prominence occurs, even with sentences that do not end with the declarative particle [da].

6) For example, in a sentence like /himajanai/ 'I am not bored,' /ja/ is predicted to have lower jaw movement than /ma/.

7) It is possible that the weakness of initial syllables may be related to the fact that it is L-toned. This is merely a speculation, though.

8) It remains a mystery at this point why the initial syllables in this sentence do not show high F1, despite their large jaw opening.

References

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

Boersma, P. and D. Weenink (1999–2015) "Praat: Doing phonetics by computer." *Software.*

Erickson, D., O. Fujimura and B. Pardo (1998) "Articulatory correlates of prosodic control, Emotion and emphasis." *Language and Speech* 41(3–4), 399–417.

Erickson, D., M. Hashi and K. Maekawa (2000) "Articulatory and acoustic correlates of prosodic contrasts: A comparative study of vowels in Japanese and English." *Journal of the Acoustical Society of Japan* 56, 265–266.

Erickson, D. (2002) "Articulation of extreme formant patterns for emphasized vowels." *Phonetica* 59(2–3), 134–149.

Erickson, D. (2004) "Perception of contrastive emphasis by American English and Japanese listeners." *Proceedings of Speech Prosody 2004*, 701–704.

Erickson, D., A. Suemitsu, Y. Shibuya and M. Tiede (2012) "Metrical structure and production of English rhythm." *Phonetica* 69(3), 180–190.

Erickson, D., S. Kawahara, J. Moore, A. Suemitsu and Y. Shibuya (2014) "Comparison of jaw displacement patterns of Japanese and American speakers of English: A preliminary report." *Onsei Kenkyu* 18(2), 88–94.

Fujimura, O. (1999) "Hatsuwa-no Kijutsuriron [Descriptive theory of utterances: C/D model]." *Journal of Acoustical Society of Japan* 55(11), 762–768.

Fujimura, O. (2000) "The C/D model and prosodic control of articulatory behavior." *Phonetica* 57(2–4), 128–138.

Fujimura, O. (2001) "Metrical vs. tonal organization of speech." In Bohumil Palek and Osamu Fujimura (eds.) Proceedings of *LP* 2000, 163–175. Prague: Charles University Press.

Fujimura, O. (2002) "Temporal organization of speech utterance: A C/D model perspective." *Cadernos de Estudos Linguisticos. Instituto de Estudos da Linguagen, Campinas* 43, 9–36.

Fujimura, O. (2003) "Stress and tone revisited: Skeletal vs. melodic and lexical vs. phrasal." In Shigeki Kaji (ed.) *Cross-Linguistic Studies of Tonal Phenomena: Histori-*

cal Development, Phonetics of Tone, and Descriptive Studies, 221–236. Tokyo: Tokyo University of Foreign Studies.

Fujimura, O. (2007) *Onseigaku Genron* [*Principles of Phonetics*]. Tokyo: Iwanami.

Fujisaki, H. and K. Hirose (1984) "Analysis of voice fundamental frequency contours for declarative sentences of Japanese." *Journal of the Acoustical Society of Japan* 5(4), 233–242.

Ishihara, S. (2003) *Intonation and Interface Conditions.* Doctoral dissertation, MIT.

Ito, J. and A. Mester (2012) *Unaccentedness in Japanese Ms.*, University of California, Santa Cruz.

Johnson, K. (2003) *Acoustic and Auditory Phonetics: 2nd Edition.* Malden and Oxford: Blackwell.

Kawahara, S. and T. Shinya (2008) "The intonation of gapping and coordination in Japanese: Evidence for Intonational Phrase and Utterance." *Phonetica* 65(1–2), 62–105.

Kawahara, S., D. Erickson, J. Moore, A. Suemitsu and Y. Shibuya (2014a) "Jaw displacement and metrical structure in Japanese: The effect of pitch accent, foot structure, and phrasal stress." *Onsei Kenkyu* 18(2), 77–87.

Kawahara, S., H. Masuda, D. Erickson, J. Moore, A. Suemitsu and Y. Shibuya (2014b) "Quantifying the effects of vowel quality and preceding consonants on jaw displacement: Japanese data." *Onsei Kenkyu* 18(2), 54–62.

Kawahara, S. (2015) "The phonology of Japanese accent." In Haruo Kubozono (ed.) *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology*, 445–491. Mouton: Mouton de Gruyter.

Keating, P. A., B. Lindblom, J. Lubker and J. Kreiman (1994) "Variability in jaw height for segments in English and Swedish VCVs." *Journal of Phonetics* 22, 407–422.

Kubozono, H. (2003) "The syllable as a unit of prosodic organization in Japanese." In Caroline Fery and Ruben van der Vijver (eds.) *The Syllable in Optimality Theory*, 99–122. Cambridge: Cambridge University Press.

Liberman, M. and A. Prince (1977) "On stress and linguistic rhythm." *Linguistic Inquiry* 8, 249–336.

Maekawa, K. (1997) "Effects of focus on duration and formant frequency in Japanese." In Y. Sagisaka, N. Campbell and N. Higuchi (eds.) *Computing Prosody*, 129–153. New York: Springer.

McCawley, J. (1978) "What is the tone language?" In Victoria Fromkin (ed.) *Tone: A Linguistic Survey*, 113–131. New York: Academic Press.

Menezes, C., B. Pardo, D. Erickson and O. Fujimura (2003) "Changes in syllable magnitude and timing due to repeated correction." *Speech Communication* 40(1–2), 71–85.

Mo, Y., J. Cole and M. Hasegawa-Johnson, (2009) "Prosodic effects on vowel production: Evidence from formant structure." *Interspeech 2009* doi=10.1.1.150.9535.

Nagahara, H. (1994) *Phonological phrasing in Japanese.* Doctoral dissertation, University of California, Los Angeles.

Pierrehumbert, J. B. and M. Beckman (1988) *Japanese Tone Structure.* Cambridge: MIT Press.

Plag, I., G. Kunter and M. Schramm (2011) "Acoustic correlates of primary and secondary stress in North American English." *Journal of Phonetics* 39(3), 362–374.

Poser, W. (1984) *The Phonetics and Phonology of Tone and Intonation in Japanese.* Doctoral dissertation, MIT.

R Development Core Team (1993–2015) "R: A language and environment for statistical computing." *R Foundation for Statistical Computing*, Vienna, Austria.

Recasens, D. (2012) "A study of jaw coarticulatory resistance and aggressiveness for Catalan consonants and vowels." *Journal of the Acoustical Society of America* 132(1), 412–420.

Schulman, R. (1989) "Articulatory dynamics of loud and normal speech." *Journal of the Acoustical Society of America* 85(1), 295–312.

Shinohara, S. (2000) "Default accentuation and foot structure in Japanese: Evidence from Japanese adaptations of French words." *Journal of East Asian Linguistics* 9(1), 55–96.

Stevens, K. (1998) *Acoustic Phonetics.* Cambridge, MA: MIT Press.

Sugahara, M. (2003) *Downtrends and Post-focus Intonation in Tokyo Japanese.* Doctoral dissertation, University of Massachusetts, Amherst.

Vatikiotis-Bateson, E. and J. A. Scott Kelso (1993) "Rhythm type and articulatory dynamics in English, French and Japanese." *Journal of Phonetics* 21, 231–265.

Williams, J. C. D. Erickson Y. Ozaki, A. Suemitsu, N. Minematsu and O. Fujimura (2013) "Neutralizing differences in jaw displacement for English vowels." *Proceedings of International Congress of Acoustics.*

(Received Feb. 15, 2015, Accepted Jun. 5, 2015)