音声研究　第18巻第２号
2014(平成26)年８月
77–87頁

Journal of the Phonetic Society
of Japan, Vol. 18 No. 2
August 2014, pp. 77–87

特集論文

Jaw Displacement and Metrical Structure in Japanese: The Effect of Pitch Accent, Foot Structure, and Phrasal Stress

Shigeto KAWAHARA*, Donna ERICKSON**, Jeff MOORE***,
Yoshiho SHIBUYA** and Atsuo SUEMITSU****

日本語の韻律特性と下顎の開き
—アクセント，韻脚，句強勢

要旨：近年の研究では，英語の単語・文レベルの強勢が下顎の開きの大きさに顕現することが示されてきた。この結果を踏まえ，本稿では日本語の韻律パターンがどのように下顎の開きに影響するかを検証した。対象としたのは日本語のピッチアクセント，韻脚（foot），韻律句（phonological phrase）の影響である。EMA を使った本実験によると，ピッチアクセントは日本語では顎の動きに影響が無く，韻脚もはっきりとした影響を及ぼしているとは言いがたい。しかし，韻律句に関しては，句の最初や最後，特に文の最後に大きな顎の開きがみられ，ここに何かしら強勢を認めることができる。ここで浮かび上がってくる仮説は,「日本語にも強勢が存在する」というもので，従来の「英語＝強勢言語，日本語＝アクセント言語」という分類法に疑問を投げかける。

Key words: jaw movements, Japanese, pitch accent, metrical foot, phrasal stress, EMA

1. Introduction

Natural languages organize their utterances into smaller rhythmical units; i.e. the observation which is often referred to as "metrical organization". The specific phonetic realizations of metrical organizations differ from one language to another. For example, English makes use of strong-weak rhythmic alternations by way of stress (e.g. Liberman and Prince 1977, Prince 1983 *et seq.*), whereas Japanese primarily makes use of pitch accent [1]. The acoustic correlates of English stress include, for example, increased duration and larger intensity (e.g. Fry 1955, 1958, Plag et al. 2011), whereas Japanese pitch accent is primarily realized as changes in F0 (Beckman 1986, Poser 1984, Sugito 1998, Weitzman 1970). The phonology of cross-linguistic patterns of metrical organization is now well understood largely owing to the development of Metrical Phonology (e.g. Liberman and Prince 1977, Gordon 2011, Hayes 1995). How English stress is realized acoustically is also well understood (e.g. Fry 1955, 1958, Gordon 2011, 2014, Plag et al. 2011).

Less well studied, however, are articulatory correlates of metrical structure in natural languages. Several recent works based on EMA (ElectroMagnetic Articulograph) show that jaw displacement may reflect metrical organization of utterances in English (Erickson et al. 2012, 2014a, b, Menezes and Erickson 2013, see also de Jong 1995, Kelso et al. 1985, Vatikiotis-Bateson and Kelso 1993): the magnitude of metrical prominence correlates well with the magnitude of jaw displacement.

Building on this research program, this paper explores a yet even less studied area of research; that is, how Japanese metrical structure is reflected in jaw displacement patterns. We specifically investigate how three aspects of Japanese metrical features—pitch accent, foot structure, and phrasal stress—manifest themselves in jaw displacement patterns.

2. Background and predictions

We start with a brief discussion of the three metrical features of Japanese (or potential lack thereof), and

* The Keio Institute of Cultural and Linguistic Studies（慶應義塾大学言語文化研究所）
** Department of General Education, Kanazawa Medical University（金沢医科大学一般教育機構）
*** Faculty of Foreign Studies, Sophia University（上智大学外国語学研究科）
**** School of Information Science, Japan Advanced Institute of Science and Technology（北陸先端科学技術大学院大学情報科学研究科）

why it is interesting to study how jaw displacement patterns interact with them.

2.1 Pitch accent in Japanese

Japanese pitch accent, unlike English stress, primarily manifests itself as a fall in F0, and effects on other phonetic dimensions such as duration or intensity are minimal, if not non-existent (Beckman 1986, Pierrehumbert and Beckman 1988, Poser 1984, Sugito 1988, Weitzman 1970). Therefore, one issue addressed in this research is how Japanese pitch accent is correlated with jaw displacement patterns; recall that jaw displacement is correlated well with English stress patterns in such a way that syllables with stronger stress show larger jaw opening (e.g. Erickson et al. 2012, 2014a, b, Kelso et al. 1985, Menezes and Erickson 2013, Vatikiotis-Baterson and Kelso 1993).

Japanese words can in principle have pitch accent on any syllable, and they can even remain unaccented (see Kawahara, in press, for a recent review). Pitch accent is phonetically characterized as an HL tonal fall, where the accented vowel receives H and the following vowel receives L—accent is represented as ' after the accented vowel. In addition, words in Tokyo Japanese receive an LH phrasal initial rise, unless the initial syllable is accented. Given these characteristics, Japanese shows a minimal triplet like (1):[2)]

(1) A minimal triplet
 a. /hashi-ga/ (LHH) “edge-NOM”
 b. /ha'shi-ga/ (HLL) “chopstick-NOM”
 c. /hashi'-ga/ (LHL) “bridge-NOM”

One characteristic of Japanese that is relevant to our study is the fact that the accented H shows a higher F0 peak than the unaccented initial H; e.g. the second syllable bearing an accentual H tone in /hashi'-ga/ “bridge-NOM” is higher than the second syllable bearing a non-accentual H tone in /hashi-ga/ “edge-NOM” (Beckman and Peirrehumbert 1988, Venditti 2005). It is not unexpected, therefore, that the presence of pitch accent may affect jaw displacement patterns in Japanese, just as in English stress. On the other hand, since Japanese pitch accent is primarily a matter of F0, it would not be surprising either if it does not affect the jaw displacement pattern.

Partly to address this question, Vatikiotis-Bateson and Kelso (1993, pp. 252–253) studied how Japanese H-L tones, some of which are associated with lexical pitch accent, affect jaw displacement, and found that H-toned syllables had smaller jaw displacement than L-toned syllables. However, they did not study the effect of accent *per se*[3)], and did not include pairs that are minimally different in terms of pitch accent in their stimulus set. This issue therefore is still a completely open, empirical question, which needs to be tested experimentally (see Fujimura 2000, 2003 for the view that pitch control and jaw displacement may be independent of one another).

2.2 Foot structure in Japanese

Feet, as metrical levels above syllables, were first proposed to account for languages that show alternate stress patterns (Hayes 1995, Liberman and Prince 1977, Selkirk 1980). It was not clear, however, until the seminal work by Poser (1990), whether pitch-accented languages like Japanese possess feet as a part of their metrical organization. Poser (1990) showed that many morphophonological word formation patterns in Japanese are based on bimoraic feet. For example, Japanese hypocoristics (diminutive name forms) are formed using a bimoraic template; e.g. from a name like /takumi/, the possible hypocoristic forms include /taku/ + /chan/, /taa/ + /chan/, and /tac/ + /chan/, but not */ta/ + /chan/. This role of feet has then been discussed in later phonological works extensively to account for several word formation patterns in Japanese (Ito 1990, Kubozono 1999, Mester 1990), as well as the distribution of pitch accent patterns (e.g. Ito and Mester 2012b, Kawahara, in press, Kubozono 1997).

There is some phonetic evidence for bimoraic feet in Japanese as well. Despite the bimoraic minimality condition mentioned above, Japanese possess lexical items that are monomoraic; e.g. /ki/ “tree”, /mi/ “content”, and /e/ “picture”. When these words appear in isolation before a phrasal boundary, they are lengthened to meet the bimoraic minimality requirement, as in [(k**i̲i̲**)(taoreta)] “The tree fell” (Braver and Kawahara 2014, Higuchi and Haraguchi 2006, Mori 2002). A recent psycholinguistic experiment by Bennett (2014) also suggests the existence of feet in Japanese metrical organization.

Besides the lengthening phenomenon discussed above, exactly how feet play a role in the phonetic implementation of Japanese is not clear. Ota et al. (2003) investigated the effect of final lengthening at the foot level in Japanese, but did not find clear evidence for it. However, they raised a possibility, though not conclusively, that feet may function as a domain of durational compensation in Japanese.

As far as we know, however, there are no studies that investigated the role of feet in the articulatory organization of Japanese. Then, it would be informative to explore if and how Japanese feet play a role in the articulatory pattern of Japanese utterances.

2.3 Phrasal stress in Japanese?

The common and dominant view of Japanese prosody is that Japanese deploys F0 to express rhythms not only at the lexical level, but also at phrasal and sentential levels (Pierrehumbert and Beckman 1988, Poser 1984, Venditti 2005; see Igarashi, in press, for a recent overview). Most researchers do not therefore posit any English-like stress in the organization of Japanese prosody, either at word-level or at phrasal level. A common description of English and Japanese is something like "English is a stress-based language, whereas Japanese is a pitch-accent language". It is often implied in this sort of statement that Japanese does not possess stress which involves some sort of strong phonetic prominence (longer duration, larger intensity, or larger jaw displacement).

One notable exception, however, is work by Fujimura (2003). Fujimura argues that Japanese does possess stress, in addition to and independent of pitch accent. In particular, he posits phrase-initial stress in his model of Japanese articulation. Many languages do show patterns of phrasal stress—some metrical prominence at the edges of phrases, and this hypothesis by Fujimura (2003) has a conceptual appeal of putting Japanese on the same footing as other languages. As Fujimura himself often puts it (p.c.), "there are no languages without stress". Whether or not Japanese possess stress is an empirical question, which should be investigated independent of what is expected from other languages. In summary, it is an interesting question whether we would see any evidence of phrasal stress in Japanese, and in this study we do so through the lens of articulatory studies.

2.4 Summary

To summarize, the current research addresses how Japanese metrical features are reflected in the articulation, focusing on the following three questions in (2):

(2) The current research questions
   a. Does pitch accent affect articulatory patterns of Japanese?
   b. Do metrical feet affect articulatory patterns in Japanese?
   c. Do we see evidence for phrasal stress in Japanese?

The current project started to investigate these questions using EMA (ElectroMagnetic Articulograph), which allows us to quantitatively track jaw displacement patterns. We report two experiments that collectively address the questions in (2).

3. Method

The current project took place in two phases: Experiment I and Experiment II. Experiment II was a follow-up of Experiment I, including all the stimuli recorded in Experiment I, in addition to some additional materials. These two experiments are reported together.

3.1 EMA

In this project 3D EMA (Carstens AG500, Japan Advanced Institute of Science and Technology (JAIST)) was used to measure the amount of jaw displacement for each syllable in the sentence. One sensor was placed between the lower medial incisors, and four additional sensors (upper incisor, nasion, left and right mastoid processes behind the ears) were used as references to correct for head movement. The occlusal plane was estimated using a biteplate with three additional sensors. The articulatory and acoustic data were digitized at sampling rates of 200 Hz and 16 kHz, respectively. In post-processing, the articulatory data were rotated to the occlusal plane and corrected for head movement using the reference sensors after low-pass filtering at 20 Hz.

The stimuli, together with stimuli for other experi-

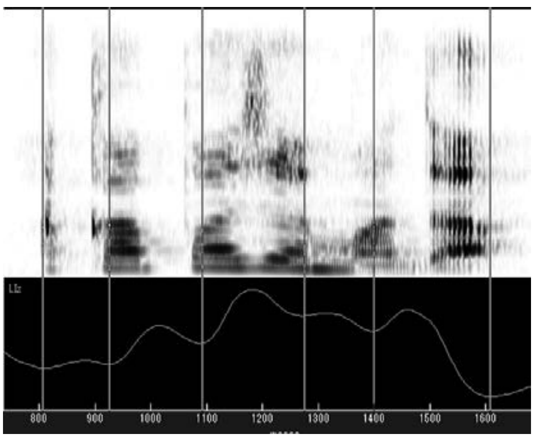

Figure 1 A sample mview from Speaker J06. The utterance is /aka-pajama da/ (ut. 8). The bottom panel shows the jaw displacement. Measurements were taken for each vowel at the time of maximum jaw opening, represented by vertical lines.

ments, were presented sentence-by-sentence in a randomized order on a PowerPoint screen in front of the speaker. The speaker pronounced each stimulus 6 times, but due to tracking errors or mispronunciation errors, sometimes number of tokens varied. Custom software (mview, Haskins Laboratories) was used to analyze the data. The lowest vertical position (maximum displacement) of the jaw[4)] with respect to the biteplane was located for each target syllable of each utterance. A sample mview screen is shown in Figure 1 to illustrate the measurement procedure. The tracing in the bottom panel shows the jaw displacement pattern. For each vowel, we measured where the jaw was most open. Maximum jaw opening sometimes occurs near the beginning of the vowel (the first three vowels), or near the end (the fourth and sixth vowel), or near the middle (the fifth vowel). The (mis)alignment of maximum jaw opening and acoustic signals is an interesting issue (see e.g. Erickson et al. 2014c), but the current research focused on the degree of maximum jaw opening, not its timing.

3.2 Stimuli

Since the magnitude of jaw displacement is non-trivially affected by vowel height (Kawahara et al. 2014, Keating et al. 1994, Menezes and Erickson 2013, Recasens 2012, Williams et al. 2013), all of the stimulus sentences were controlled for vowel quality. In order to examine the effect of pitch accent on jaw displacement (and lack thereof), the stimulus set included the triplet (inspired by Fujimura 2003) given in (3). Recall that ' represents the presence of pitch accent on the preceding vowel.

(3) Accent triplet
  a. hashi-ga aru
     edge-NOM exists
     “There is an edge.”
  b. ha'shi-ga aru
     chopstick-NOM exists
     “There is a chopstick.”
  c. hashi'-ga aru
     bridge-NOM exists
     “There is a bridge.”

In order to examine the effect of phrasal stress, the stimulus set also included sentences having only syllables with /a/:

(4) All /a/ sentences
  a. aka-gasa da
     red-umbrella COPULA
     “That’s a red umbrella.”
  b. aka-pa'jama da
     red-pajima COP
     “That’s a red pajama.”
  c. da'kara ma'na wa atama'-ga
     that’s why Mana-TOPIC hair-NOM
     sarasara da
     smooth COP
     “That’s why Mana’s hair is smooth”

In addition to those stimuli used in Experiment I ((3) and (4)), additional stimuli were prepared for Experiment II. In particular, we added sentences with /a/, which do not end with the declarative /da/ particle[5)].

(5) All /a/ sentences
  a. da'kara ma'na wa atama'ga
     that’s why Mana-TOP head
     kata'katta
     hard-PAST
     “That’s why Mana was so stubborn.”
  b. da'kara ma'na-wa atama'ga
     that’s why Mana-TOP head
     kata'katta ka?
     hard-PAST QUESTION
     “Is that why Mana was so stubborn?”

3.3 Speakers

For Experiment I, one female speaker of Japanese in her 20’s participated. Although she grew up in Hakata, Kyushu, she speaks Tokyo Japanese in her daily life without people generally noticing a Kyushu accent. For Experiment II, two native speakers of Japanese, both in their 30’s, participated: one male (J07), a native speaker of Tokyo Japanese and one female (J08), a native speaker of Kanazawa dialect[6)]. All speakers produced the standard Tokyo accentual patterns for the *hashi* triplet[7)]. In terms of their English proficiency, J06 was an advanced intermediate, J07, an advanced bilingual, and J08, a poor speaker of English.

4. Results

4.1 The accent triplet

Starting with the effect of accent on jaw displacement, Figures 2–4 (J06, J07, J08) show the jaw displacement patterns of the minimal *hashi*-triplet that differs from one another only in terms of accents. The y-axis represents the amount of jaw displacement in mm, measured as a distance from the placement of

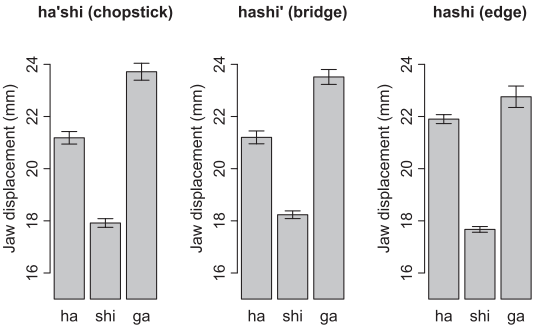

Figure 2 The Jaw displacement patterns of the accentually minimal triplet by Speaker J06. Y-axis represents the degree of jaw displacement (mm). The error bars represent standard errors. To show observed jaw displacement patterns, different figures in this paper use different y-axis scale.

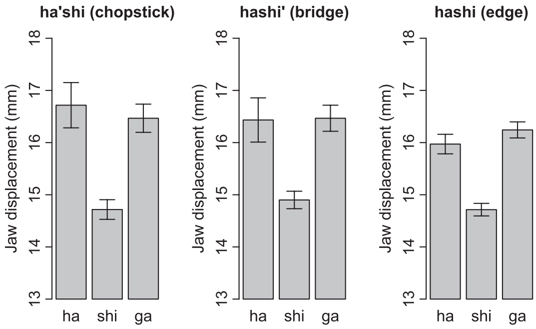

Figure 4 The jaw displacement patterns of the accent triplet, Speaker J08.

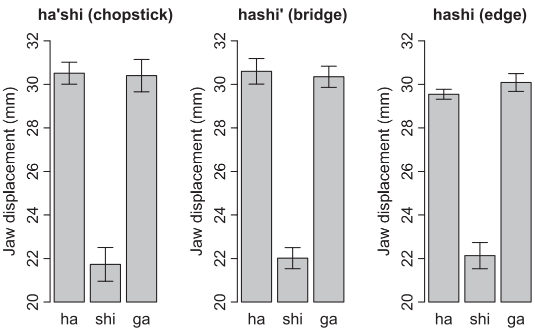

Figure 3 The jaw displacement patterns of the accent triplet, Speaker J07.

lower incisor and the bite plane. In this figure and others throughout this paper, the error bars represent the standard errors of the means [8].

We observe that for none of the speakers are there any substantial differences among the three sub-figures. The first syllables in the leftmost sub-figures are accented, but they are not substantially larger than the first syllables in the other two figures for these speakers. For Speaker J06, in fact, the first syllable shows the largest opening in the rightmost sub-figure.

As for Speaker J08 (Figure 4), the initial syllable is slightly higher in the left most panel (accented) than the other two, and the middle syllable is slightly higher in the mid panel (accented) than the other two. To test whether these small differences are substantial, jaw displacement was regressed with accentedness, vowel quality, and their interactions as independent variables. The result shows that only vowel quality had a significant effect ($t = -7.6, p < .001$), not accent ($t = 1.5, n.s.$) or the interaction ($t = -0.5, n.s.$). It seems safe to conclude that accent patterns *per se* do not affect jaw displacement pattern.

The small jaw displacement of the second syllable with respect to the first and third syllables in these figures may come from the fact that the vowel is a closed, high vowel /i/, compared to the other two vowels /a/ (Erickson et al. 2013, Kawahara et al. 2014, Menezes and Erickson 2013, Williams et al. 2013). It may alternatively be that initial and final syllables receive some sort of prominence, a possibility that we address with the next set of data.

4.2 Metrical structure and phrasal stress

The following is a discussion of metrical structure and phrasal stress based on analysis of sentences (4) and (5), sentences with all /a/.

Figures 5–7 show jaw displacement patterns of /aka-gasa da/ and /aka-pajama da/. All speakers show the largest jaw opening in initial and final syllables for both sentences. It seems likely that this large jaw opening represents the existence of initial and final stress in Japanese (as argued by Fujimura 2003) [9]. For possible evidence of a phrase-final prominence from Japanese speakers' L2 speech, see Erickson et al. (2014b).

Regarding the patterning of the non-final syllables, one idea to model their patterns is to say that initial syllables are highest, and they decline gradually, just like F0 declination (Fujisaki and Hirose 1984, Pierrehumbert and Beckman 1988, Poser 1984). An alternative idea that we would like to entertain is that these syllables (moras) are grouped into bimoraic feet, within which a strong-weak relation is assigned. By assigning a strong-weak relationship between feet, we can model the observed pattern. The Liberman and Prince (1977) style of

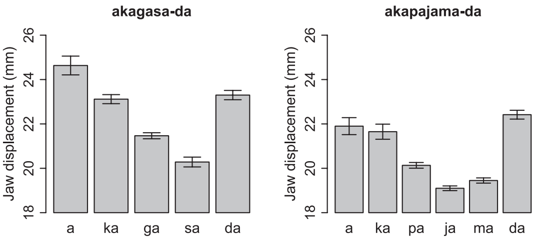

Figure 5 The jaw displacement patterns of /aka-gasa da/ and /aka-pajama da/, Speaker J06.

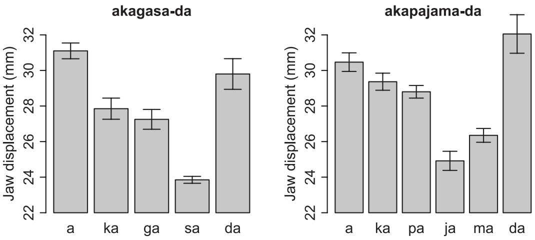

Figure 6 The jaw displacement patterns of /aka-gasa da/ and /aka-pajama da/, Speaker J07.

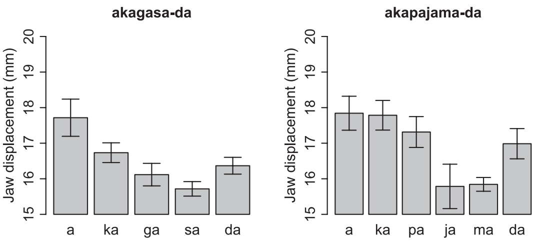

Figure 7 The jaw displacement patterns of /aka-gasa da/ and /aka-pajama da/, Speaker J08.

prosodic phrasing (see also Prince 1980, Selkirk 1980), as shown in Figure 8, captures these patterns well[10].

One virtue of this approach is that it allows us to account for the slightly larger opening of /ma/ in the /aka-pajama da/ sentence with respect to the preceding /ja/, by positing a metrical organization illustrated in

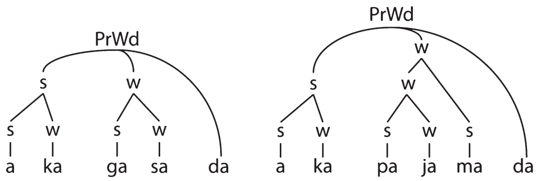

Figure 8 Possible foot-based analyses of the jaw displacement patterns in Figures 5, 6 and 7. "s" = strong; "w" = weak. /da/ is the strongest of all, perhaps receiving special phrasal stress, and therefore it remains unmarked.

Figure 8(b)—/ma/ may be metrically stronger than /ja/[11]. Admittedly this evidence for the effect of feet on jaw displacement is post-hoc and does not offer a conclusive argument, but we would like to suggest this possibility for future discussion.

Figures 9–11 show the jaw displacement patterns of long sentences consisting of /a/ (recall that Speaker J06 produced only one sentence of this type). These sentences are less straightforward to interpret than the other two simple sentences. We start with the first two speakers, Speakers J06 and J07.

Overall, both of the speakers behave similarly in that sentence-final jaw opening is larger than phrase-final jaw opening (no matter what the final syllable), suggesting that Japanese possesses sentence-final stress.

For the sake of analysis, let us posit the following phrasing[12] in (9) for Speakers J06 and J07, which makes sense from the syntactic perspective: the first phrase contains topic phrases; the second phrase contains the nominative phrases; the final phrase contains the predicates.

(9)

a. (dakara mana wa)(atama ga)(sarasara da)
b. (dakara mana wa)(atama ga)(katakatta)
c. (dakara mana wa)(atama ga)(katakatta ka)

Given these phrasings, we observe that both speakers show final prominence at the end of each phrase,

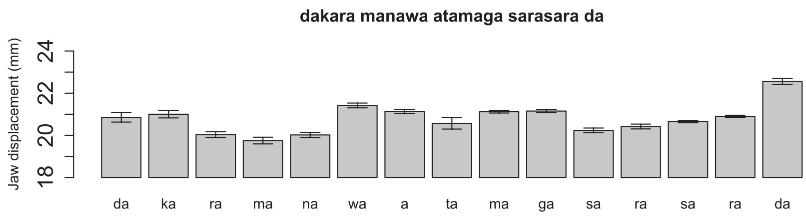

Figure 9 The jaw displacement pattern of /dakara mana-wa atama-ga sarasara-da/, Speaker J06.

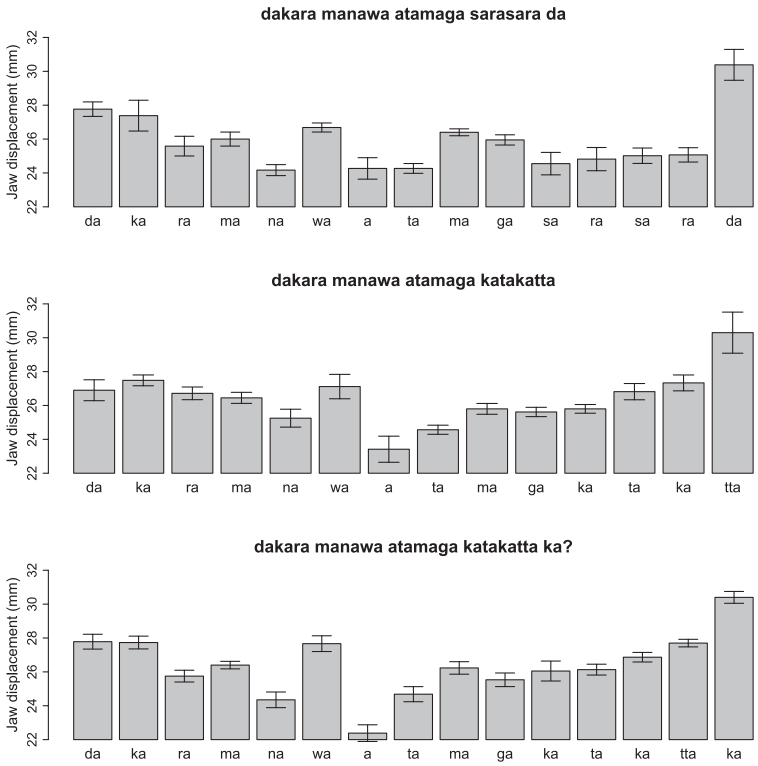

Figure 10 The jaw displacement patterns of the long sentences consisting of /a/, Speaker J07.

although the sentence-final stress is stronger. Concretely, we observe some prominence on /wa/, /ga/, and /da, ta, ka/. Among these three syllables, the sentence-final prominence on /da, ta, ka/ is largest, suggesting that Japanese possess phrase-final stress, as well as sentence-final stress, the latter of which is stronger.

However, unlike those in Figures 5–7, they do not show clear phrase-initial jaw opening, except in sentence-initial position: in the second and third phrases, none of the initial syllables show clear large jaw opening. Also for these speakers, the first two sentence-initial syllables (i.e., /daka/) show large jaw opening. It may be the case that there is sentence-initial stress, but not phrase-initial stress.

Now let us consider Speaker J08, who shows slightly different patterns. As shown in Figure 11, she shows consistent large jaw opening on /ra/ at the end of /dakara/. This speaker may have phrased /dakara/ separately from /mana-wa/. Also, she does not show large opening at the end of the nominative phrase /atama ga/. It is possible that Speaker J08 groups the nominative phrase and the predicate in the same phonological phrase, as in (10).

(10)

a. (dakara)(mana wa)(atamaga sarasara da)
b. (dakara)(mana wa)(atamaga katakatta)
c. (dakara)(mana wa)(atamaga katakatta ka)

We acknowledge that it is dangerous to speculate on phrasal patterning for each speaker based on jaw displacement patterns, and argue for the effect of phrasing on jaw displacement patterns at the same time. Nevertheless, what is consistent across the three speakers is large opening on those vowels that end a syntactic phrase, hence, likely a phonological phrase (/daka**ra**/, /mana-**wa**/ and sentence-final particles), rather than elsewhere. We do not observe prominence in any syllables that are in the middle of syntactic/phonological phrasing (e.g. /na/ in /manawa/ or /ta/ in /atama/).

To summarize, in all sentences that consist of /a/ we observed phrase-final large prominence, in particular, in sentence-final position. We tender this as evidence that Japanese possesses phrase-final and sentence-final stress. All the sentences recorded in this experiment also show evidence for sentence-initial stress, but perhaps not phrase-initial stress.

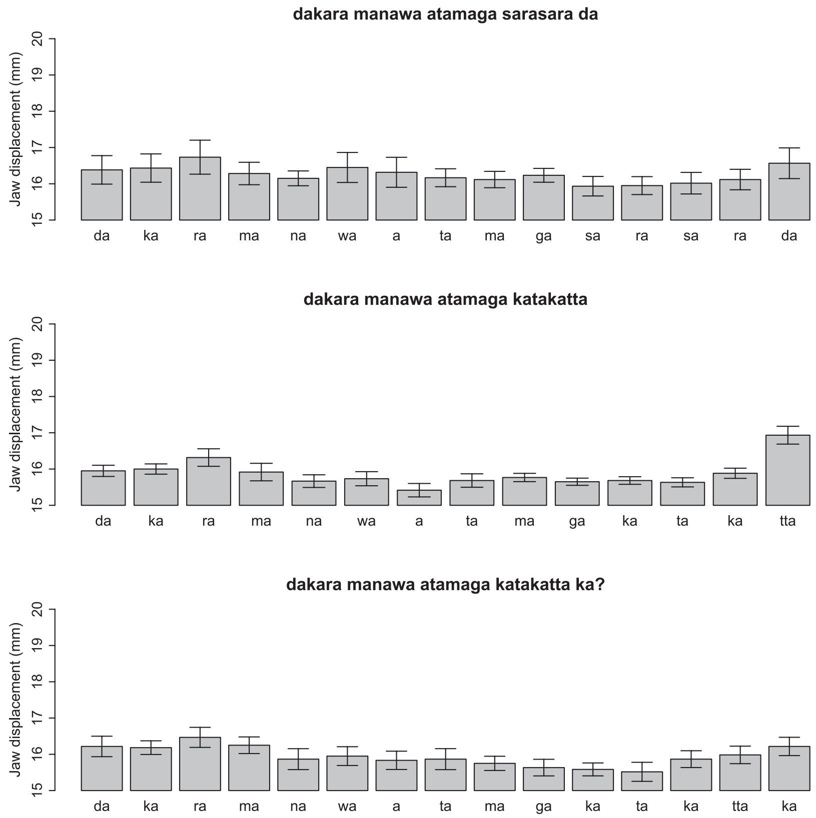

Figure 11 The jaw displacement patterns of the long sentences consisting of /a/, Speaker J08.

One question raised by John Kingston (p.c.) regarding the current conclusion is whether this "stress" shows similar phonetic correlates as contrastive focus, as contrastive focus realizes by way of strong prominence, e.g., phrasal stress in languages like English. We find this to be a very interesting question. Erickson et al. (2000) report an experiment that addresses this issue. They found that (i) Japanese jaw displacement is not exaggerated as much as English, but that (ii) /a/ and /e/ do show magnified jaw displacement when focused, whereas /i/ does not. These patterns seem to be compatible with what we observe in the current study, and hope to investigate this issue further.

5. General discussion

Our current study is exploratory, as there has not been much work on the relationship between metrical structure and jaw displacement in Japanese. Therefore, our conclusions remain tentative.

However, despite the limited number of speakers who participated in this study, we found some patterns that are consistent across speakers. First, we observed little if any effects of pitch accent on jaw displacement. This result may not be a surprise because Japanese pitch accent is primarily a matter of F0; Fujimura (2003) also explicitly modeled the implementation of F0 patterns and jaw movement patterns separately. At any rate, we have found an intriguing cross-linguistic difference: whereas English stress is known to affect jaw displacement patterns, Japanese pitch accent does not (for the effect of English stress on jaw displacement, see Erickson et al. 2012, 2014a, b, Kelso et al. 1985, Menezes and Erickson 2013, Vatikiotis-Bateson and Kelso 1993, de Jong 1995).

Also, for (short) sentences with /a/, we observed some sort of "declination" in which jaw opening starts large and decreases gradually. We entertained a possibility that this apparent declination pattern may arise from the foot structure of Japanese, although we need to remain tentative about this hypothesis. There is a simple, equally plausible declination analysis as well.

Finally, the current experiments seem to find good evidence for phrase-final and sentence-final stress in Japanese, in that we constantly observe large jaw opening in the final syllables; the evidence for phrase-initial prominence was less consistent but nevertheless observed in some sentences. In sentences that consist of multiple phrases, we found large jaw opening at the end of each phrase, and the sentence-final stress is more prominent than sentence-internal prominence. We found these final prominence patterns throughout our results, regardless of the type of sentence-final particles. Japanese, then, is possibly a language with stress, as envisioned by Fujimura; "there are no languages without stress".

Acknowledgments

This research is supported by the Japan Society for the Promotion of Science, Grants-in-Aid for Scientific Research #26770147 and Keio Gijuku Academic Development Fund to the first author and #22520412 to the second author. Portions of Experiment I were presented at ICPP 3 at NINJAL. We would like to thank Haruka Fukazawa, John Kingston, Hinako Masuda, Shin-ichiro Sano, Yoko Sugioka, Yukiko Sugiyama, and two anonymous reviewers for their comments on this research. We also thank Professor Jianwu Dang for the use of the EMA laboratory at JAIST. All remaining errors are ours.

Notes

1) Sometimes within a single language (like English), the term "stress" is used to refer to word-level prominence, whereas the term "pitch accent" is used to refer to phrase- or utterance-level prominence (see Gordon 2014 for an overview). This paper uses "stress" for the prominence observed both at word- and phrasal-level (in English and Japanese), and uses "pitch accent" to refer to Japanese lexical accent.

2) For the sake of exposition, this paper uses *romaji*-transcription instead of IPA transcriptions. We use phenomic // marks to represent sounds, except when phonetic details are relevant.

3) It seems that they did not distinguish between H*L accent which are associated with lexical accent and other phrasal accent (e.g. initial LH). They say "in this study, tone level rather than accent is treated as the relevant binary variable analogous to the stress distinctions used for French and English" (p. 235). Our experiment differs from this strategy in that it directly tests the effect of pitch accent by making a triplet that differs only in pitch accent.

4) By measuring the distance from the occlusal plane to maximum amount of jaw opening in the syllable, we account for differences in palate shape across different speakers. Also, we do not average jaw displacement across speakers, since we are interested in differences across the speakers.

5) Also included were sentences with syllables /e/, /o/ and /i/. (We were unable to create natural Japanese sentences consisting of only syllables with /u/.) However, these sentences are not discussed here due to space limitation.

6) J07 is the first author of this paper, and this is partly due to the fact that it is difficult to find a speaker who is willing to participate in an EMA experiment. While we admit that this use of our own speech is not ideal, we also assume that we cannot consciously control details of jaw displacements. Again we interpret J07's results cautiously, and invite the readers to replicate our results.

7) We checked against the pitch contours, although we cannot show them in this paper due to space limitation.

8) We feel it is important to continue to use bargraphs, instead of line graphs, to display the results, since we are working within the framework of the C/D model (Fujimura (2000), see also Erickson et al. (2012) and other related works), in which patterns of jaw displacement represent syllable pulse patterns, reflecting patterns of syllable magnitude in an utterance.

9) This sentence-final jaw-opening is observed in Figure 2 of Vatikiotis-Bateson and Kelso (1993). Their data, which shows reiterant speech on /ma/ for the sentence /obaasan wa momo o hirotte, ie ni motte kaerimashita/, is consistent with our data.

10) The recursive footing of /pajama/ in Figure 5(b) is not what is usually assumed in the phonological literature; it is often assumed that /paja/ are footed, and /ma/ is directly attached to the PrWd (Ito and Mester 1992/2004). On the other hand, it does not seem to us that positing this recursive footing is too radical either. However, we also reiterate that our proposal about using feet to explain the pattern in Figure 8(b) is tentative.

11) An alternative explanation emerged from another research of ours (Kawahara et al. 2014): coronals inhibit more jaw displacement than labials, and obstruents inhibit more jaw displacement than sonorants (see also Vatikiotis-Bateson and Kelso 1993: 240). Given this, the larger jaw displacement of /ma/ with respect to /ja/ may come from these consonantal effects.

12) There is a question about exactly which phrase of metrical organization that this "phrase" refers to. Since there are a number of different, competing theories on prosodic organization in Japanese, we would like to remain neutral about it, although we suggest that we are talking above a phrase that is right above Word; i.e. Minor Phrase (McCawley 1968), accentual phrase (Pierrehumbert and Beckman 1988, Venditti 2005), or simply, (phonological) phrase (Ito and Mester 2012a).

References

Beckman, M. (1986) *Stress and Non-Stress Accent*. Dordrecht: Foris.

Bennett, R. (2012) *Foot-conditioned phonotactics and prosodic constituency*. Doctoral dissertation, University of California at Santa Cruz.

Braver, A. and S. Kawahara (2014) "Incomplete vowel lengthening: A first study." In R. Santana-LaBarge, (ed.) *Proceedings of WCCFL 31*. Somerville: Cascadilla Press.

de Jong, K. (1995) "The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation." *Journal of the Acoustical Society of America* 97, 491–504.

Erickson, D., M. Hashi and K. Maekawa (2000) "Articulatory and acoustic correlates of prosodic contrasts: A comparative study of vowels in Japanese and English." *J. Acoust. Soc. Jp* 56, 265–266.

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

Erickson, D., S. Kawahara, J. Moore, A. Suemitsu and Y. Shibuya (2013) "Effect of vowel height position on jaw opening patterns in Japanese." Talk presented at *International Conference on Phonetics and Phonology (ICPP) 3*, NINJAL: Tokyo.

Erickson, D., S. Kawahara, Y. Shibuya, A. Suemitsu and M. Tiede (2014a) "Comparison of jaw displacement patterns of Japanese and American speakers of English: A preliminary report." *Journal of the Phonetic Society of Japan* 18(2), 88–94.

Erickson, D., S. Kawahara, J. C. Williams, J. Moore, A. Suemitsu and Y. Shibuya (2014b) "Metrical structure and jaw displacement: An exploration." *Speech Prosody 2014*, 300–303.

Erickson, D., S. Kawahara, J. Moore, C. Menezes, A. Suemitsu, J. Kim and Y. Shibuya (2014c) "Calculating articulatory syllable duration and phrase boundaries." *International Symposium on Speech Production 2014*, 102–105.

Fry, D. B. (1955) "Duration and intensity as physical correlates of linguistic stress." *Journal of the Acoustical Society of America* 27, 1765–1768.

Fry, D. B. (1958) "Experiments in the perception of stress." *Language and Speech* 1, 126–152.

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

Fujimura, O. (2003) "Stress and tone revisited: Skeletal vs. melodic and lexical vs. phrasal." In S. Kaji (ed.) *Cross-linguistic studies of tonal phenomena: Historical development, phonetics of tone, and descriptive studies*, 221–236. Tokyo: Tokyo University of Foreign Studies.

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

Gordon, M. (2011) "Stress: Phonotactic and phonetic evidence." In M. van Oostendorp, C. J. Ewen, E. Hume and K. Rice (eds.) *The Blackwell companion to phonology*, 924–948. Oxford: Wiley-Blackwell.

Gordon, M. (2014) "Disentangling stress and pitch accent: Toward a typology of prominence at different prosodic levels." In H. van der Hulst (ed.) *Word stress: theoeretical and typological issues*, 83–118. Oxford: Oxford University Press.

Hayes, B. (1995) *Metrical Stress Theory: Principles and Case Studies*. Chicago: The University of Chicago Press.

Higuchi, M. and S. Haraguchi (2006) "Final lengthening in Japanese." On-in Kenkyu [Phonological Studies] 9, 9–16.

Igarashi, Y. (in press) "Intonation." In H. Kubozono, (ed.) *The handbook of Japanese language and linguistics: Phonetics and phonology*. Berlin: Mouton.

Ito, J. (1990) "Prosodic minimality in Japanese." CLS 26-II: *Papers from the Parasession on the Syllable in Phonetics and Phonology*, 213–239.

Ito, J. and A. Mester (2012a) "Recursive prosodic phrasing in Japanese." In T. Borowsky, S. Kawahara, T. Shinya and M. Sugahara (eds.) *Prosody matters*, 280–303. London: Equinox Publishing.

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

Kawahara, S. (in press) "The phonology of Japanese accent." In H. Kubozono (ed.) *The handbook of Japanese language and linguistics: Phonetics and phonology*. Mouton: Mouton de Gruyter.

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

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.

Kelso, J. A. S., E. Vatikiotis-Bateson, E. L. Saltzman and B. Kay (1985) "A qualitative dynamic analysis of reiterant speech production: phrase portraits, kinematics, and dynamic modeling." *Journal of the Acoustical Society of America* 77, 266–280.

Kubozono, H. (1997) "Lexical markedness and variation: A nonderivational account of Japanese compound accent." In B. Agbayani and S.-W. Tang (eds.) *Proceedings of West Coast Conference on Formal Linguistics* 15, 273–287. CSLI Publications.

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

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

McCawley, J. D. (1968) *The Phonological Component of a Grammar of Japanese*. The Hague: Mouton.

Menezes, C. and D. Erickson (2013) "Intrinsic variations in jaw deviation in English vowels." *Proceedings of International Congress of Acoustics*, 060253.

Mester, A. (1990) "Patterns of truncation." *Linguistic Inquiry* 21, 475–485.

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

Ota, M., D. R. Ladd and M. Tsuchiya (2003) "Effects of foot structure on mora duration in Japanese?" *Proceedings of the 15th International Conference on Phonetic Sciences*, 459–462.

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, 362–374.

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

Poser, W. (1990) "Evidence for foot structure in Japanese." *Language* 66, 78–105.

Prince, A. (1980) "A metrical theory for Estonian quantity." *Linguistic Inquiry* 11, 511–562.

Prince, A. (1983) "Relating to the grid." *Linguistic Inquiry* 14, 19–100.

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

Selkirk, E. (1980) "The role of prosodic categories in English word stress." *Linguistic Inquiry* 11, 563–605.

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

Sugito, M. (1998) *Nihongo onsei no kenkyu 5: "hana" to "hana"* [Studies on Japanese Phonetics 5]. Osaka: Izumi Shoin.

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

Venditti, J. (2005) "The ToBI model of Japanese intonation." In S.-A. Jun (ed.) *Prosodic typology: The phonology of intonation and phrasing*, 172–200. Oxford: Oxford University Press.

Weitzman, R. S. (1970) *Word accent in Japanese*. Detroit, MI: Management Information Services.

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*, 060268.

(Received Jun. 15, 2014, Accepted Sep. 16, 2014)