Acoust. Sci. & Tech.

TECHNICAL REPORT

The phonetics of jaw displacement in Japanese vowels
Jaw displacement in Japanese vowels

Shigeto Kawahara,[1] Donna Erickson[2] and Atsuo Suemitsu[3]*

Keio University[1]
Haskins Laboratory[2]
Sapporo University of Health Sciences[3]

Abstract: The jaw is one of the most important articulators in speech production. Despite this, we know next to nothing about how Japanese speakers use their jaw to produce vowels. Against this background, in order to explore the articulatory natures of Japanese vocalic jaw movements, this paper presents a detailed, quantitative EMA study of the five vowels in Japanese, focusing on the following four specific questions: (1) How many mms does the jaw open for each type of vowel in Japanese?; (2) Does the presence of an onset consonant affect the degrees of jaw opening?; (3) Does the speed of the jaw movement vary depending on how much the jaw opens?; (4) What is the reliable acoustic correlate of the jaw opening? In answer to these questions, the current experiment demonstrates that (1) In Japanese, the degree of the jaw opening is in the order of [a] > [e] > [o] > [i] > [u]; (2) The presence of onset consonant [p] generally decreases the jaw opening; (3) The degrees of the jaw opening and its speed positively correlate; and (4) F1 and duration are reliable acoustic correlates of the jaw opening. Implications of these results for phonetic theories are discussed throughout the paper.

Keywords: Vowel, jaw, displacement, speed, acoustics

PACS number: 43.70.Aj, 43.70.Bk, 43.70.Fq, 43.70.Kv

1. Introduction

In the history of phonetic research on speech articulation, more attention tends to be paid on tongue movement than on jaw movement. However, a number of researchers have shown that the jaw is nevertheless one important articulator in speech production. It has been shown, for example, that the jaw opens wider for stressed vowels than for unstressed vowels in English; contrastive emphasis likewise results in larger jaw opening in English, Japanese and other languages [1–12]. This body of research shows that speakers make use of the jaw to express metrical and semantic prominence. There are also studies on effects of metrical phrasing on articulatory kinematics in terms of jaw movement [13–15], which found systematic patterns. Some previous work shows that speakers use the jaw to express sentential rhythm in such a way that a vowel that is more prominent within a sentence shows a larger jaw opening [7, 16–18]. These findings led to a theory of speech articulation, such as Fujimura's C/D model [19–21], in which the jaw is designated as the primary articulator, through which phonological representations are mapped on to phonetic gestures.

At the segmental level, differences in vowel height—or sometimes called vowel "openness" [22–24]—result in different degrees of the jaw openness [25–29]. We believe that this is a highly understudied area, in general but particularly in Japanese. Most if not all phonetics textbooks, when explaining vowel articulations, focus on tongue gestures, not mentioning any jaw movement patterns (e.g. [30–32], among others). When it comes to Japanese, we know next to nothing about how the jaw moves when articulating vowels. The only quantitative, articulatory study on the jaw opening in Japanese vowels that we are aware of is [33], which is preliminary in the sense that the work analyzed only two speakers. Extending on this work, we report more comprehensive data on the jaw movements of the five different vowels in Japanese. Against the theoretical background laid out above, this paper explores the phonetics of the jaw movement in five Japanese vowels with the following specific questions in mind. (For a study addressing the effect of contrastive emphasis on Japanese jaw movement, see [3]; for studies address the effects of Japanese phrasal prominence and lexical accent, see [34, 35]—we do not deal with these topics in this paper.)

(1) **How many mms does the jaw open for each type of vowel in Japanese?** The answer to this ques-

* Corresponding email address: `kawahara@icl.kei.ac.jp`.

tion has descriptive value in and of itself: we want to know precisely how much the jaw opens for each different vowel, just as, say, we want to know the formant characteristics of each vowel. Getting these concrete measures on jaw opening is of some significance, partly because vowel height is sometimes characterized as "openness" (e.g. in the current IPA system). Quantitative measures are also useful for studies on the jaw movement using sentences with different vowel qualities; by knowing how much vowel quality affects the jaw opening, we can potentially "wash away" these effects, using the algorithm proposed by [28, 29]. This analysis allows researchers to study the impact of prosody on jaw movement, without worrying about the effects of vowel quality.

(2) **Does the presence of an onset consonant affect the jaw opening in Japanese?** This question is related to the general issue of coarticulation in phonetic research. How does the consonant and the vowel interact in terms of the jaw movement? Coarticulation is well-studied for tongue movement as well as for laryngeal gestures (see [36]). How about the jaw? (see e.g. [27]). [33] showed effects of consonant-vowel coarticulation with coronal consonants such as [t] and [n] (but again, with only two speakers). On the other hand, a study on English speakers using video recordings [37] did not show strong evidence for consonant-vowel articulation for jaw movement. Given the results of these previous studies, it would be interesting to test whether labial consonants coarticulate with vowels, when they use independent articulators (cf. [38,39]). It is also interesting to test whether coarticulation in jaw movement is observed across different speakers in general.

(3) **Does the speed of the jaw movement vary depending on how much the jaw opens in Japanese?** We can conceive of two hypotheses about this question: (i) the speed is constant across different vowels vs. (ii) those vowels that involve larger jaw opening involves higher speed; behind this second hypothesis lies a principle that speakers make vowel duration more constant than otherwise. This hypothesis is implied in an oft-heard statement like "low vowels are longer because the jaw has to travel longer" (e.g. [40]) (though see [41–43] for counterevidence of this claim). Some other work [13, 44] on the other hand, presents evidence for the latter hypothesis that the speed and displacement degree may correlate with each other. To the extent that Japanese is mora-timed (see [45] for a critical overview), speakers may attempt to keep the durations of different vowels constant to the extent possible, predicting different speed profiles for different vowels. To the best of our knowledge, this question has not been addressed for Japanese speech.

(4) **What is the acoustic correlate of the jaw opening?** Previous studies have shown that the primary acoustic correlate of vowel openness is F1 in English [16] (see also [46,47]), where larger jaw opening corresponds to higher F1. This correlation holds presumably because the Helmholtz resonation, which is responsible for F1, positively correlates with the area function of the Helmholtz resonator, as given in the equation (1) (adapted from [48]). A_c, which is the numerator of the equation (1), is larger for more open vowels, resulting in higher F1. (For previous acoustic analyses of Japanese vowels, which deal with other acoustic features including higher formant values, see [49–53]).

$$f = \frac{c}{2\pi}\sqrt{\frac{A_c}{A_b l_b l_c}} \qquad (1)$$

(where c is the speed of sound, A is the area function, l is the length, $_c$ is the constricted tube, $_b$ is the tube behind the constriction)

There are other acoustic cues that are known to sometimes correlate with vowel openness. For example, F0 is often reported to be higher for higher—or more closed—vowels [50, 54–56]. Intensity has been thought to be an acoustic correlate of sonority, where more open vowels are more sonorous [57, 58]. [57] cites many studies which have found correlation between sonority and intensity level, some of which include vowels of different openness (p.111). The experiment reported in [57], however, did not find this correlation between intensity and vowel openness either in English or Spanish. This inconsistency may come from the fact that on the one hand, the more open the mouth is, the less energy is trapped in the mouth (see [59]); on the other hand, lower F1—characteristics of closed vowels—results in higher intensity [47]. F0 and intensity seem to show less consistent correlation with jaw opening. Finally, the more open vowel is usually longer [40, 60–62]. However, a recent study [63] who studied the vowel duration patterns in the Corpus of Spontaneous Japanese (the CSJ: [64]) found that there is very large variability, so much so that durational differences among vowels may not be that reliable in real speech.

Against these previous studies, we test F0, F1, intensity, and duration to investigate how robustly each of these acoustic measures correlate with jaw openness in the production of Japanese speech.

2. Method

2.1. EMA

This study made use of 3D EMA (ElectroMagnetic Articulograph), which allows us to obtain quantitative measurements of the articulatory movements. The current experiment used Carstens AG500, hosted at Japan Advanced Institute of Science and Technology (JAIST), Ishikawa, Japan. The crucial target sensor was placed on the lower incisor of the speakers. Four additional sensors were attached to the speakers in order to correct for the head movement during recording. The distance between the bite plane and the lower incisor was calculated, as a quantitative measure of the jaw opening. See Fig 1 for illustration.

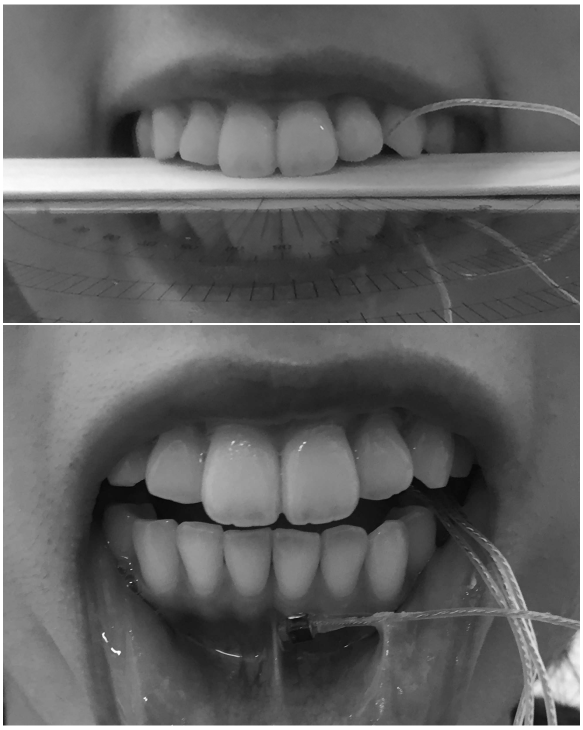

Fig. 1 Illustration of the estimation of the bite plane (top) and the placement of a sensor on the lower incisor (bottom). These pictures were taken for a different experiment, but they are close enough to illustrate the current experiment.

2.2. Speakers

The participants were five native speakers of Japanese, coded as SpS (Male), SpL (Female), SpH (Female), SpT (Female), and SpY (Female) in this paper. All speakers were in their 20's or 30's at the time of recording.

2.3. Procedure

The current study analyzes the following data: All the five vowels in Japanese—[a], [i], [u], [e], [o], pronounced in syllables without onset consonants, and those with initial [p] (i.e. [pa], [pi], [pu], [pe], [po]). We analyze those syllables with [p] for a few reasons. First, since [p] uses the lips, it would be interesting to examine whether it affects the jaw movement, as the lips and the jaw are in principle independent articulators. Finding interaction between [p] and the following vowel would be robust evidence for coarticulation between vowels and preceding consonants (recall that [33] already found evidence for consonant-vowel coarticulation with consonants produced with the tongue). Second, using syllables with [p] had a virtue of, during the measurement phase, being able to isolate the consonantal gesture from the vocalic gesture.

To control for the order effect, our stimulus structure took the form of a Latin-Square matrix, shown in Table 1.

Table 1 The stimulus structure

[(p)a]	[(p)i]	[(p)u]	[(p)e]	[(p)o]	[(p)a]
[(p)i]	[(p)u]	[(p)e]	[(p)o]	[(p)a]	[(p)i]
[(p)u]	[(p)e]	[(p)o]	[(p)a]	[(p)i]	[(p)u]
[(p)e]	[(p)o]	[(p)a]	[(p)i]	[(p)u]	[(p)e]
[(p)o]	[(p)a]	[(p)i]	[(p)u]	[(p)e]	[(p)o]

Since there are five vowels in Japanese, the stimulus structure had five rows. The order of the first row ([a]-[i]-[u]-[e]-[o]) is the standard way of reciting vowels in Japanese orthography. Within each row, all the five vowels appeared, and the initial vowel was repeated at the end of each line to avoid the effect of phrase-final large jaw opening in Japanese [34, 35] (cf. [65] for general phrase-edge strengthening effects). However, upon the examination of the obtained data, no effects of the phrase-final opening were observed; therefore, all data were included in the analysis. In this structure, all the vowels appear 6 times, and the position in which each vowel occurs is controlled.

All the speakers repeated this matrix 5 times, which was presented via a Powerpoint screen, mixed with many other stimuli of other experiments.

2.4. EMA analysis

The analysis of the EMA data was conducted using mview, custom developed at Haskins Laboratory. An example of the mview screen is given in Fig 2. For the displacement data, using the lp_snapex function (an mview-internal function), the maximum point of jaw displacement was automatically detected for each vowel, both for onsetless syllables and those with [p]. The jaw speed was measured with syllables with onset [p]. This is because the speakers "reset" the jaw opening with these syllables, as they need to close their lips for [p]—for onsetless syllables, locating velocity maximum was hard to locate. Maximum velocity of the jaw was measured at the onset of each syllable, which represents its max velocity in the opening phase for all the target syllables. The lp_peaks function in mview is used for this measurement.

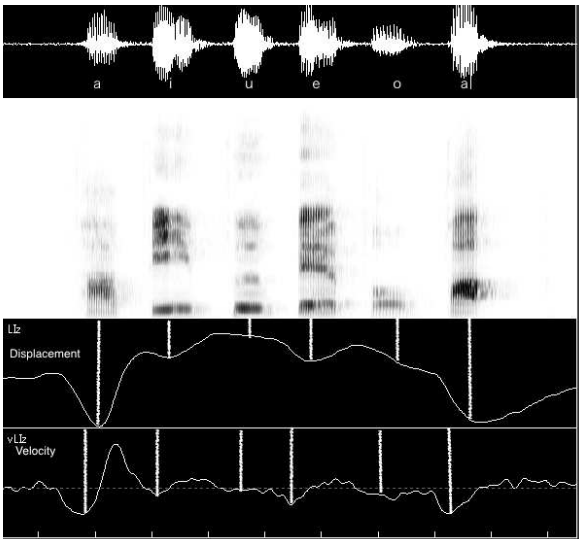

Fig. 2 An example illustrative figure of mview screen for EMA measurement.

2.5. Acoustic analysis

The speech signals were recorded simultaneously with the EMA recording, at a sampling rate of 16 kHz. The acoustic analysis is based on the analysis of onsetless syllables, in order to avoid coarticulatory effects of [p]. Despite the instructions to the contrary, Speakers SpL and SpH did not put a pause between each vowel, and therefore acoustic measurement was impossible for them, because boundaries between each vowel were too blurry to locate precisely. Speaker SpY forgot to put a pause for about half of the tokens. For this speaker, only measurable tokens were used in the following analysis (108 tokens out of 180).

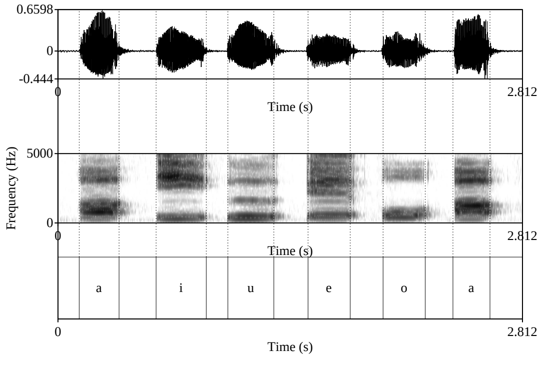

Fig. 3 Illustration of the acoustic labeling using Praat.

Using Praat [66, 67], boundaries for each vowel were placed using the annotation function, as shown in Fig. 3. A script was written to automatically detect the midpoint of each vowel based on these annotations, and to create a 20 ms analysis window centering that midpoint. Average values for F1, F0, intensity and duration within these analysis windows were then extracted.

3. Results

3.1. EMA results

Table 2 shows degrees of the jaw displacement for all the speakers. Here and throughout, the values in parentheses show margins of errors for 95% confidence intervals. We observe that all the speakers show a clear tripartite distinction based on vowel height: [a] is open, both [e] and [o] are semi-open, and [i] and [u] are closed. This result closely corresponds to the characterization of these vowels in the current IPA system. (Perhaps the only exception is SpL, whose difference between [a] and [e] is not as big as other speakers.) This finding is good evidence that one articulatory correlate of "vowel height" is the jaw opening [25–28]. Second, within each pair of the same height, the front vowels ([e] and [i]) are more open than the corresponding back vowels ([o] and [u]), although this tendency is very small for SpS.

Table 2 The amount of jaw displacement (mm) in syllables without onset. The values in parentheses show margins of errors for 95% confidence intervals.

	[a]	[e]	[o]	[i]	[u]
SpS	32.8 (0.8)	30.2 (0.6)	29.9 (0.4)	25.7 (0.6)	25.2 (0.5)
SpL	18.5 (0.2)	18.2 (0.2)	17.5 (0.2)	16.8 (0.2)	16.3 (0.2)
SpH	30.5 (0.9)	26.8 (0.6)	25.5 (0.6)	20.2 (0.4)	19.3 (0.3)
SpT	23.3 (0.6)	21.9 (0.5)	20.9 (0.6)	18.2 (0.3)	17.2 (0.4)
SpY	31.3 (1.1)	27.8 (0.7)	26.0 (0.7)	22.1 (0.8)	20.7 (0.2)

We note in passing that in English too, within a pair of the same height, front vowels tend to show larger jaw opening than back vowels [28,29], as shown in Fig 4. The similarity between Japanese and English we found here is interesting, and should be explored in other languages in future research.

Table 3 shows the jaw displacement degrees in syllables with onset [p]. Overall, compared to Table 2, the jaw opens less; this pattern holds for all the vowels for all the speakers, except for the two high vowels for SpT. The general ordering among the five vowels—[a] > [e] > [o] > [i] > [u]—still holds in these syllables, although there are some reversals (e.g. [o] and [i] for SpS). Overall, therefore, [p] decreases the jaw opening in general, but "inherent degrees" of the jaw opening for different vowels remain the same. This new finding adds to

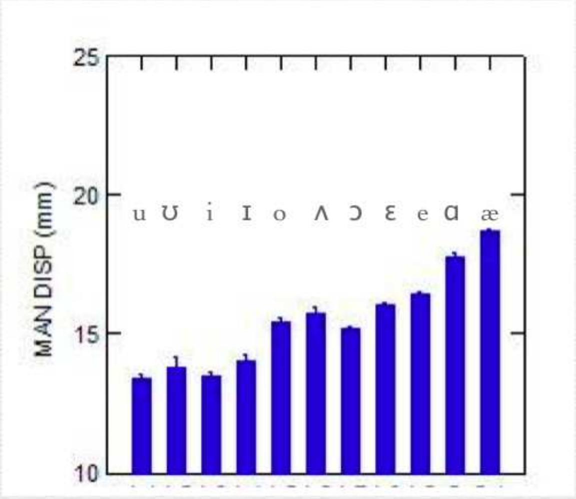

Fig. 4 The degrees of jaw opening for English vowels. Based on [28, 29].

what [33] showed with coronal consonants. It seems that consonant-vowel coarticulation is prevalent in jaw displacement patterns in Japanese (cf. [37]).

Table 3 The jaw displacement (mm) in syllables with onset [p].

	[a]	[e]	[o]	[i]	[u]
SpS	26.8 (1.2)	26.7 (1.3)	22.8 (0.9)	24.4 (0.8)	21.8 (0.5)
SpL	17.7 (0.3)	16.9 (0.2)	17.0 (0.2)	16.5 (0.2)	16.0 (0.1)
SpH	23.4 (0.4)	21.5 (0.2)	21.7 (0.4)	19.9 (0.2)	18.7 (0.2)
SpT	21.5 (0.3)	20.1 (0.2)	20.1 (0.3)	18.7 (0.2)	18.2 (0.2)
SpY	25.3 (0.2)	24.3 (0.3)	24.1 (0.3)	21.8 (0.1)	21.1 (0.2)

Table 4 shows the maximum speed of the jaw with syllables that have onset [p]. As with the degrees of openness, the observed order is generally [a] > [e] > [o] > [i] > [u] (although we do see some reversals; e.g. [a] vs. [e] and [o] vs. [i] for SpT). In other words, the more open the vowel, the faster the jaw moves [13]. This is not compatible with the idea that the speed of the jaw is kept constant across different vowels, and instead supports the proposal that speakers keep "articulatory duration" more or less constant.

To visualize correlation between degrees of the jaw displacement and speed, Fig. 5 shows a scatterplot between these two measurements. We observe that for a vowel that involves larger opening, the faster the jaw moves. We could take this to be evidence that Japanese speakers may possibly control the duration constant (at least more constant than otherwise), just like English speakers studied in [13] (cf. the observed addressed under the notion of "stiffness" in task dynamics; see e.g.

Table 4 The maximum speed of the jaw for each vowel (cm/sec) in syllables with onset [p].

	[a]	[e]	[o]	[i]	[u]
SpS	4.2 (0.6)	1.9 (0.4)	0.6 (0.2)	1.0 (0.2)	0.5 (0.3)
SpL	1.5 (0.2)	1.5 (0.2)	0.6 (0.2)	0.3 (0.10)	0.2 (0.1)
SpH	6.8 (0.6)	6.9 (0.5)	4.8 (0.4)	2.3 (0.4)	0.8 (0.1)
SpT	4.4 (0.3)	4.6 (0.3)	1.5 (0.3)	1.9 (0.2)	0.5 (0.1)
SpY	6.5 (0.5)	8.3 (0.6)	3.0 (0.6)	1.1 (0.2)	0.2 (0.1)

[44]). This result suggests that the oft-made statement that "low vowels are longer because the aw has to travel longer" is not necessarily true (see also [41, 42]). This sort of statement assumes that the speed of the jaw is constant across different vowels, which is proven not to be true in the current results.

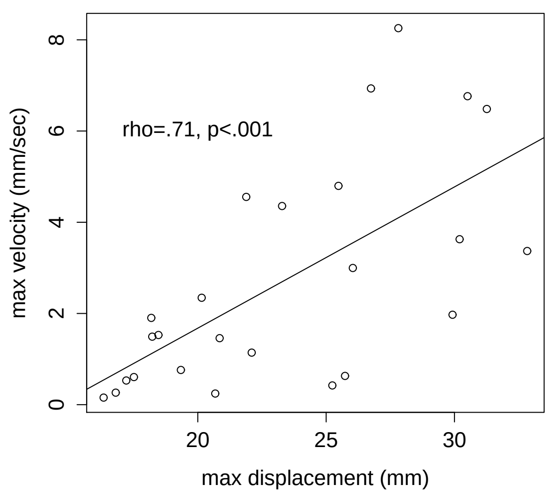

Fig. 5 Correlation between degrees of the jaw displacement degree and maximum speed.

Of course, the correlation is not perfect, and therefore it is not the case that all vowels have the exact same duration, as we will confirm in Table 8 later (see also [40, 60–62] for work addressing durational differences between vowels). Duration is more constant across different vowels than if speakers were using the same speed for all the vowels.

3.2. Acoustics

Tables 5, 6, 7, 8 show F1, F0, intensity and duration of each of the five vowels recorded in the current study. Starting with Table 5, F1 correlates with the jaw opening in such a way that the more open vowels have higher F1.

One interesting reversal found between the jaw open-

Table 5 F1 (Hz).

	[a]	[e]	[o]	[i]	[u]
SpS	861 (21)	457 (13)	454 (13)	302 (8)	321 (5)
SpT	792 (54)	401 (17)	407 (20)	333 (14)	368 (12)
SpY	881 (31)	468 (17)	441 (18)	354 (19)	390 (13)

ing and F1 is between [i] and [u]. Despite the fact that [i] involves more jaw opening than [u] (see Tables 2 and 3), F1 is lower for [i] than [u]. We conjecture that this "reversal" occurs because tongue backness affects Helmholtz resonation as well. Recall the equation given in (1), where the Helmholtz resonation is lowered as the length of the back cavity gets longer [48]. The length of the back cavity is longer for [i] than [u], because [i] is a front vowel, resulting in lower F1. These results show that the jaw opening crucially affects F1, but is not the only factor that determines F1. Front vowels show lower F1 than back vowels, because of how Helmholtz resonation is determined. These characteristics of the Helmholtz resonation explains why [e] and [i] show lower—or comparable—F1 than [o] and [u], despite the former being more open.

Concerning F0 and intensity (Tables 6 and 7), we do not seem to observe systematic covariation as a function of vowel openness. In fact, in the current results, F0 and intensity do not seem to differ very much between the five different vowels.

Table 6 F0 (Hz).

	[a]	[e]	[o]	[i]	[u]
SpS	114 (6.9)	113 (7.7)	113 (8.7)	115 (7.0)	118 (8.0)
SpT	212 (8)	211 (8)	214 (7)	203 (16)	218 (10)
SpY	200 (13.5)	199 (14.5)	207 (8.6)	187 (19.3)	213 (10.2)

Table 7 Intensity (dB).

	[a]	[e]	[o]	[i]	[u]
SpS	76 (1.0)	75 (0.9)	75 (0.8)	75 (0.9)	75 (1.0)
SpT	84 (1.2)	85 (1.0)	84 (0.9)	85 (1.0)	86 (0.7)
SpY	77 (1.3)	76 (1.1)	78 (1.5)	76 (1.1)	78 (1.0)

As for duration, Table 8 shows that [a], which tends to involve large jaw opening, tends to be longer. High vowels—[i] and [u]—on the other hand tend to be shorter, and recall that they show small jaw opening. These observations suggest that duration does correlate with jaw opening, despite speakers' effort to move the jaw quicker for a vowel that involves large jaw opening.

Table 8 Duration (ms).

	[a]	[e]	[o]	[i]	[u]
SpS	192 (10)	194 (9)	193 (12)	191 (9)	181 (8)
SpT	199 (12)	198 (12)	202 (9)	195 (9)	192 (10)
SpY	241 (19)	237 (19)	228 (17)	236 (17)	224 (17)

To visually examine the correlations between jaw openness on the one hand and F1, F0, intensity, and duration on the other, Fig. 6 provides three scatterplots. To avoid obtaining spurious correlations due to inter-speaker differences, these values are normalized within each speaker. It shows that F1 and duration, and not the other two acoustic measures, correlates with the jaw opening.

This result shows that F1 is a fairly robust acoustic correlate of jaw opening (ρ=0.86), supporting the proposal put forward by [16]. This correlation makes sense given how F1 is affected by the area function of the constriction [46–48] (though see the discussion above on the effect of tongue backness on F1). Duration too correlates significantly with jaw opening, despite the fact that the jaw moves faster for a vowel that requires large jaw opening. Thus, although speakers may attempt to keep articulatory duration for each vowel constant, the perfect isochrony is not achieved, at least not acoustically.

Another intriguing aspect of this result is that it shows that intensity does not necessarily correlate with the jaw opening (cf. [59])—just because the jaw is more open, it is not necessarily the case that more energy comes out of the mouth (see also [58]).

4. Conclusion

The current study is situated in the emerging research tradition, which has shown that the jaw is one of the important articulators for speech production. This research tradition has often focused on how metrical or semantic prominence impacts the degree of jaw opening. Less well studied, in our opinion, is the effect of vowel quality on the degrees of jaw opening, and precisely how the jaw is used in producing these five vowels in Japanese. We aimed to fill that gap.

In that spirit, this study addressed the following questions: (1) How many mms does the jaw open for each type of vowel in Japanese?; (2) Does the presence of an onset consonant affect the jaw opening?; (3) Does the speed of the jaw movement vary depending on how much the jaw opens?;(4) What is the acoustic correlate of the jaw opening?

In answer to these questions, the current experiment, based on EMA and acoustic data, demonstrates the following. (1) In Japanese, the degree of the jaw opening is in the order of [a] > [e] > [o] > [i] > [u]. The more "open" (or lower) vowels indeed show larger jaw opening, and within a pair of the same height, the jaw opens more for front vowels. (2) The presence of onset consonant [p] decreases the jaw opening, which shows that the jaw and the lips coarticulate in Japanese. This expands on a previous, preliminary, finding about the coarticulation between consonants and vowels by [33] (cf. [37]). (3) The degrees of the jaw opening and its speed positively correlate, which can be taken as evidence that speakers may attempt to make the "articulatory duration" constant, if not perfectly, across different vowels [13]. (4) F1 and duration are reliable acoustic correlates of jaw opening, at least more reliable than F0 or intensity [16].

We hope that this study provides a first step toward understanding how Japanese speakers—and speakers of other languages, for that matter—use the jaw to produce speech sounds.

Acknowledgements

Thanks to Osamu Fujimura and C.J. Williams for their guidance. We also thank Professor Jianwu Dang, for the use of his lab at JAIST. This work is supported by JSPS, Grants-in-Aid for Scientific Research #26770147 and #26284059 to the first author and #22520412 and #25370444 to the second and third authors.

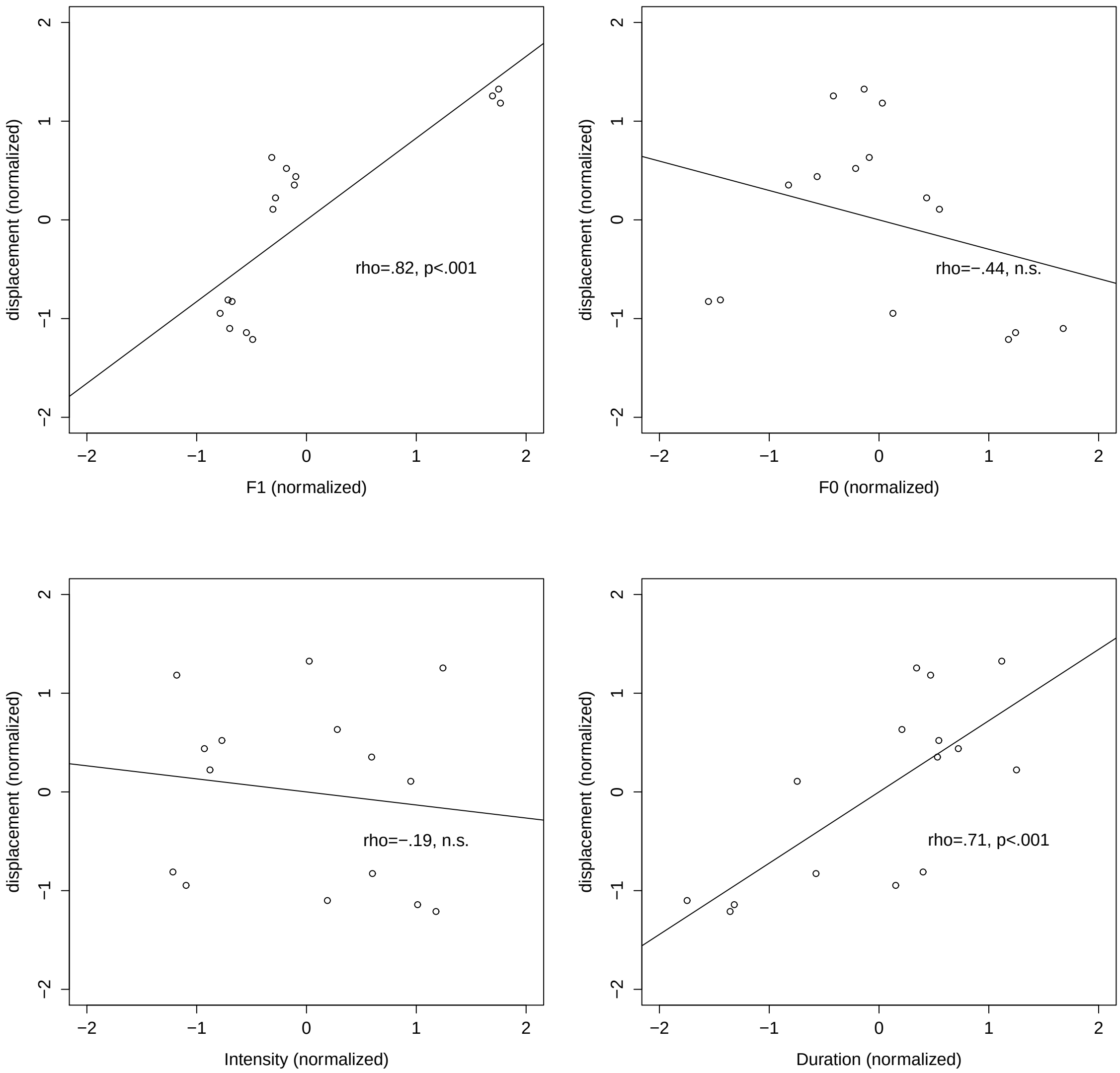

Fig. 6 Correlations between the jaw displacement degrees and four acoustic features (F1, F0, intensity, duration). All values were normalized for each speaker.

REFERENCES

[1] Beckman ME, Edwards J. Articulatory evidence for differentiating stress categories. In: Keating PA, editor. Papers in laboratory phonology III: Phonological structure and phonetic form. Cambridge, UK: Cambridge University Press; 1994. p. 7–33.

[2] de Jong K. The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. Journal of the Acoustical Society of America. 1995;97:491–504.

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

[4] Stone M. Evidence for a rhythm pattern in speech production: Observations of jaw movement. Journal of Phonetics. 1981;9:109–120.

[5] Erickson D. Effects of contrastive emphasis on jaw opening. Phonetica. 1998;55:147–169.

[6] Erickson D. Articulation of extreme formant patterns for emphasized vowels. Phonetica. 2002;59:134–149.

[7] Erickson D, Kawahara S. Articulatory correlates of metrical structure: Studying jaw displacement patterns. Linguistic Vanguard. 2016;.

[8] Harrington J, Fletcher J, Beckman M. Manner and place conflicts in the articulation of accent in Australian English. In: Broe M, Pierrehumbert JB, editors. Papers in Laboratory Phonology V: Acquisition and the Lexicon. Cambridge: Cambridge University Press; 2000. p. 40–51.

[9] Kent RD, Netsell R. Effects of stress contrasts on certain articulatory parameters. Phonetica. 1971;24:23–44.

[10] Menezes C, Pardo B, Erickson D, Fujimura O. Changes in syllable magnitude and timing due to repeated correction. Speech Communication. 2003;40:71–85.

[11] Macchi M. Segmental and suprasegmental features and lip and jaw articulations [Doctoral dissertation]. New York University; 1985.

[12] Maekawa K, Kiritoni S, Imagawa H, Honda K. Effects of focus on mandible movement. Journal of Acoustical Society of Japan. 1998; p. 259–260.

[13] Edwards J, Beckman ME, Fletcher J. Articulatory kinematics of final lengthening. Journal of the Acoustical Society of America. 1991;89:369–382.

[14] Edwards J, Beckman M. Articulatory timing and the prosodic interpretation of the syllable. Phonetica. 1988;45:156–174.

[15] Nittrouer S, Munhall KG, Kelso JAS, Tuller B, Harris KS. Patterns of interarticulator phrasing and their relation to linguistic structure. Journal of the Acoustical Society of America. 1998;84(5):1653–1661.

[16] Erickson D, Suemitsu A, Shibuya Y, Tiede M. Metrical structure and production of English rhythm. Phonetica. 2012;69:180–190.

[17] Erickson D, Kawahara S, Williams CJ, Moore J, Suemitsu A, Shibuya Y. Metrical structure and jaw displacement: An exploration. Speech and Prosody 2014. 2014;.

[18] Vatikiotis-Bateson E, Kelso JAS. Rhythm type and articulatory dynamics in English, French and Japanese. Journal of Phonetics. 1993;21:231–265.

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

[20] Fujimura O. Stress and tone revisited: Skeletal vs. melodic and lexical vs. phrasal. In: Kaji S, editor. Cross-Linguistic Studies of Tonal Phenomena: Historical Development, Phonetics of Tone, and Descriptive Studies. Tokyo: Tokyo University of Foreign Studies; 2003. p. 221–236.

[21] Fujimura O, Williams CJ. Remarks on the C/D model. Journal of the Phonetic Society of Japan. 2015;19(2):2–8.

[22] Chomsky N, Halle M. The Sound Pattern of English. New York: Harper and Row; 1968.

[23] Clements GN. Vowel height assimilation in Bantu languages. In: Hubbard K, editor. Proceedings of Berkeley Linguistic Society 17S: Proceedings of the Special Session on African Language Structures. Berkeley: Berkeley Linguistic Society; 1991. p. 25–64.

[24] Clements N, Hume E. The internal organization of speech sounds. In: Goldsmith JA, editor. The Handbook of Phonological Theory. Cambridge and Oxford: Blackwell; 1995. p. 245–306.

[25] Keating PA, Lindblom B, Lubker J, Kreiman J. Variability in jaw height for segments in English and Swedish VCVs. Journal of Phonetics. 1994;22:407–422.

[26] Lee SH. Orals, gutturals, and the jaw. In: Connell B, Arvaniti A, editors. Papers in Laboratory Phonology IV: Phonology and Phonetic Evidence. Cambrdige University Press; 1994. p. 343–360.

[27] Recasens D. A study of jaw coarticulatory resistance and aggressiveness for Catalan consonants and vowels. Journal of the Acoustical Society of America. 2012;132(1):412–420.

[28] Williams JC, Erickson D, Ozaki Y, Suemitsu A, Minematsu N, Fujimura O. Neutralizing differences in jaw displacement for English vowels. Proceedings of International Congress of Acoustics. 2013;.

[29] Menezes C, Erickson D. Intrinsic variations in jaw deviation in English vowels. Proceedings of International Congress of Acoustics. 2013;.

[30] Someya M. Basic English pronunicatin for Japanese. Sanshusha; 2000.

[31] Reetz H, Jongman A. Phonetics. Oxford: Blackwell-Wiley; 2008.

[32] Ladefoged P. Vowels and consonants. Blackwell; 2001.

[33] Kawahara S, Masuda H, Erickson D, Moore J, Suemitsu A, Shibuya Y. Quantifying the effects of vowel quality and preceding consonants on jaw displacement: Japanese data. Onsei Kenkyu [Journal of the Phonetic Society of Japan]. 2014;18(2):54–62.

[34] Kawahara S, Erickson D, Moore J, Suemitsu A, Shibuya Y. Jaw displacement and metrical structure in Japanese: The effect of pitch accent, foot structure, and phrasal stress. Onsei Kenkyu [Journal of the Phonetic Society of Japan]. 2014;18(2):77–87.

[35] Kawahara S, Erickson D, Suemitsu A. Edge prominence and declination in Japanese jaw dislpacement patterns: A view from the C/D model. Journal of the Phonetic Society of Japan. 2015;19(2):33–43.

[36] Hardcastle WJ, Hewett N, editors. Coarticulation: theory, data, and techniques. Cambridge: Cambridge University Press; 1999.

[37] Wilson I, Erickson D. Effect of syllable onset, coda, and nucleus on degree of skin stretching over the mandible. Proc of International Congress of Acoustics. 2013;19:doi:10.1121/1.4799467.

[38] Iskarous K. Locus equations are an acoustic express of articulatory synergy. Journal of the Acoustical Society of America. 2010;128:2021–2032.

[39] Sussman H, McCaffrey H, Matthews S. An investigation of locus equations as a source of relational invariance for stop place categorization. Journal of the Acoustical Society of America. 1991;90:1309–1325.

[40] Lehiste I. Suprasegmentals. Cambridge: MIT Press; 1970.

[41] Nooteboom SG, Slis I. A note on the degree of opening and the duration of vowels in normal and “pipe” speech. IPO annual progress report. 1970;5:55–58.

[42] Lisker L. On “explaining” vowel duration variation. Tech Report. 1974;SR-37/38.

[43] Toivonen I, Blumenfeld L, Gormley A, Hoiting L, Logan J, Ramlakhan, et al. Vowel height and duration. Proceedings WCCFL 32. 2015;.

[44] Munhall KG, Ostry DJ, Parush A. Characteristics of velocity profiles of speech movements. Journal of Experimental Psychology: Human Perception and Performance. 1985;11(4):457–474.

[45] Warner N, Arai T. Japanese Mora-Timing: A Review. Phonetica. 2001;58:1–25.

[46] Lindblom B, Sundberg J. Acoustical consequences of lip, tongue, jaw, and laynx movement. Journal of the Acoustical Society of America. 1971;62:693–707.

[47] Stevens K. Acoustic Phonetics. Cambridge, MA: MIT Press; 1998.

[48] Johnson K. Acoustic and Auditory Phonetics: 2nd Edition. Malden and Oxford: Blackwell; 2003.

[49] Chiba T, Kajiyama M. The Vowel: Its Nature and Structure. Tokyo: Tokyo-Kaiseikan Publishing Co.; 1941.

[50] Hirahara T, Kato H. The effect of F0 on vowel identification. In: Tohkura Y, Vatikiotis-Bateson EV, Sagisaka Y, editors. Speech Perception, Production and Linguistic Structure. Tokyo: Ohmsha; 1992. p. 89–112.

[51] Hirahara T, Akahane-Yamada R. Acoustic analysis of Japanese vowels. Proceedings of ICA. 2004;(3287-3290).

[52] Keating PA, Huffman M. Vowel variation in Japanese. Phonetica. 1984;41:191–207.

[53] Nishi K, Strange W, Akahane-Yamada R, Kubo R, Trent-Brown SA. Acoustic and perceptual similarity of Japanese and American English vowels. Journal of the Acoustical Society of America. 2008;124(1):576–588.

[54] Diehl R, Kluender K. On the objects of speech perception. Ecological Psychology. 1989;1:121–144.

[55] Honda K, Fujimura O. Intrinsic vowel F0 and phrase-final lowering: Phonological vs biological explanations. In: Gauffin J, Hammerberg B, editors. Vocal Fold Physiology: Acoustic, Perceptual, and Physiological Aspects of Voice Mechanisms. San Diego: Singular Publishing Group; 1991. p. 149–157.

[56] Maddieson I. Phonetic universals. In: Hardcastle W, Laver J, editors. The Handbook of Phonetics Sciences. Oxford: Blackwell; 1997. p. 619–639.

[57] Parker S. Quantifying the Sonority Hierarchy [Doctoral Dissertation]. University of Massachusetts, Amherst; 2002.

[58] Parker S. Sound level protrusions as physical correlates of sonority. Journal of Phonetics. 2008;36:55–90.

[59] Schulman R. Articulatory dynamics of loud and normal speech. Journal of the Acoustical Society of America. 1989;85:295–312.

[60] Campbell N. Segmental elasticity and timing in Japanese. In: Tohkura Y, Vatikiotis-Bateson EV, Sagisaka Y, editors. Speech Perception, Production and Linguistic Structure. Ohmsha; 1992. p. 403–418.

[61] Keating PA. Universal phonetics and the organization of grammars. In: Fromkin VA, editor. Phonetic Linguistic Essays in Honor of Peter Ladefoged. Orlando: Academic Press; 1985. p. 115–132.

[62] Sagisaka Y, Tohkura Y. Kisoku-niyoru onsei goosei-no tame-no onin jikan seigyo [Phoneme duration control for speech synthesis by rule]. Denshi Tsuushin Gakkai Ronbunshi. 1984;67:629–636.

[63] Kawahara S, Shaw J. Effects of vowel informativity on vowel duration in Japanese;. Ms. Keio University.

[64] NINJAL. Nihongo hanashikotoba koopasu [The Corpus of Spontaneous Japanese], Ver. 2.; 2008. Database.

[65] Fougeron C, Keating P. Articulatory strengthening at edges of prosodic domain. Journal of the Acoustical Society of America. 1997;106:3728–3740.

[66] Boersma P. Praat, a system for doing phonetics by computer. Glot International. 2001;5(9/10):341–345.

[67] Boersma P, Weenink D. Praat: Doing phonetics by computer; 1999–2016. Software.

Author's profile To be inserted if accepted.

Cover Letter for Acoustical Science and Technology

(1) **Title of paper**
The phonetics of jaw displacement in Japanese vowels
Jaw displacement in Japanese vowels

(2) **Full name(s) of author(s)**
Shigeto Kawahara,[1] Donna Erickson[2] and Atsuo Suemitsu[3*]

(3) **Affiliation(s)**
Keio University[1]
Haskins Laboratory[2]
Sapporo University of Health Sciences[3]

(4) **Approximately five keywords**
Vowel, jaw, displacement, speed, acoustics

(5) **PACS number**
43.70.Aj, 43.70.Bk, 43.70.Fq, 43.70.Kv

(6) **Short running title**
Jaw displacement in Japanese vowels

(7) **Category of article:**
TECHNICAL REPORT

(8) **Mailing address**
2-15-45 Mita, Minato-ku Tokyo, JAPAN 108-8345

(9) **Classification**
Speech

(10) **Number of pages**
TEXT: 10 FIGURES: 6 TABLES: 8