On the Pronunciation of Low Vowels in Japanese Newscaster's Speech

Philip Spaelti
Department of English, Kobe Shoin Women's University
spaelti[at]shoin.ac.jp

Abstract

An informal observation made over many years, leads me to believe that Japanese newscasters produce phrase final vowel [a] in a consistently lowered and fronted manner. Using samples from NHK news, an analysis was performed which would seem to indicate that this is indeed the case. Further investigation seems warranted.

Keywords: Japanese vowels, formants, spectral analysis

1. Introduction

Japanese has one of the more straightforward vowel systems: A simple five-vowel system, which is perhaps the single most common type of vowel system around the world. The features most remarked upon are the lack of rounding on the high back vowel [u]—which is generally realized as [ɯ]—as well as the devoicing of high vowels in voiceless contexts. In this short note I would like to draw attention to something that to my knowledge at least has not been remarked on, namely the realization of low vowels, most notably in the speech of newscasters.

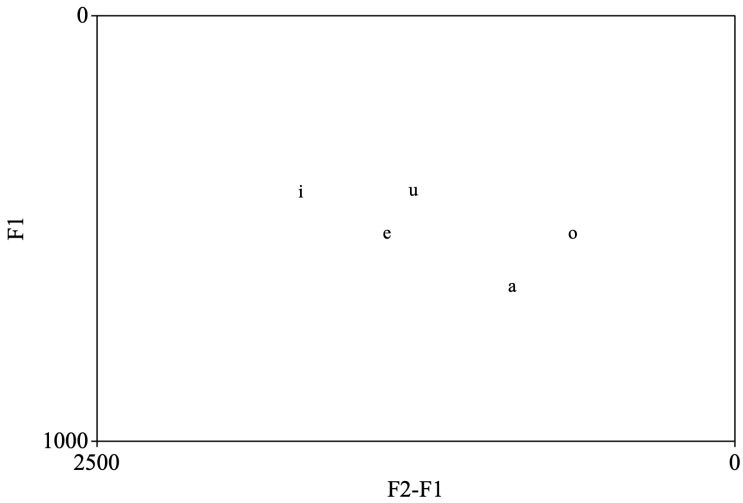
The observation described in this paper is that the low vowel [a] of Japanese shows contextually determined variation in its realization in the speech of newscasters. Specifically [a] in phrase final position is realized in a characteristically lowered or fronted manner, making it perceptually more similar to English [æ], and distinct from the more central or back realization of [a] more generally. The primary evidence for this is my own personal observation. In my experience this has been a consistent feature of newscasters most notably on the Japanese national broadcaster NHK for many years.

And while this will not be a primary claim of this paper, I have similarly observed that the same fronting does not seem to occur in the speech of others speakers on television or indeed among Japanese speakers generally. In the research described here the focus will be on a preliminary investigation to determine whether my informal observations are supported by the evidence. Further investigation will be needed to determine how widespread this variation is, and whether it is truly particular to newscasters speech exclusively.

2. The Japanese vowel system

That the Japanese vowel system is a five-vowel system is a well-established fact (Ladefoged & Maddieson 1996; Vance 2008). This is also made obvious by the language's kana writing system. The standard five-vowel system, represented by languages such as Spanish, consists of the three poles of the vowel triangle, a—i—u, with an [e] interpolated between [a] and [i] and an [o] between [a] and [u]. The one point on which the Japanese system deviates from the textbook example is [u], which is generally produced with little or no lip-rounding. This results in a quality for the [u] that makes it much less distinct from [i] then would be expected. This can be seen in the diagram (1), created from the samples discussed in more detail in the following sections.

(1)



Diagrams such as this show the distribution of the vowels in terms of their formants F1 and F2, with F1 on the vertical axis increasing from top to bottom, and F2 on the horizontal increasing from right to left. This type of diagram allows for a straightforward comparison of the vowels in terms of their acoustics, which can easily be measured using computer software, with traditional descriptions in terms of their articulation. The articulatory interpretation of the diagram is that the vertical dimension represents vowel height, actualized through jaw opening and the amount of tongue lowering. Meanwhile the horizontal dimension represents vowel fronting, i.e., the combination of forward motion of the tongue and lip-rounding.

In diagram (1) we can readily see the effect of the unrounded pronunciation of [u] which leads to this vowel appearing much further to the left than might be expected.

The simple system of Japanese can be contrasted with that of English, which typically contrasts 11 or more vowel qualities. In particular, instead of the single vowel [a] at the bottom, English has several contrasting vowels; typically examples such as *hat, hot*, and *hut*. In such a system the realization of each of these vowels will of necessity need to be quite precise in order to maintain the contrast to the other members of the set. On the other hand, a language such as Japanese, which has only one vowel in the lower area, may allow for a much broader realization of the relevant vowel.

The claim being made in this paper is that the low vowel is realized in a varied fashion, but that the variation is not random, but is instead correlated with the position in the utterance;

that [a], when it appears at the end of a phrase, has a realization distinct from when it appears elsewhere in the speech of Japanese newscasters.

To investigate this, samples were collected from Japanese TV news and analyzed.

3. Procedure

Short samples were taken from Japanese TV News. All samples came from NHK BS news. A total of 4 samples were collected, spoken by 3 different announcers; two male, and one female.

The samples were transferred to the computer using Handbrake, and the audio track was extracted using Audacity.

The text of the samples was transcribed to regular Japanese orthography with the assistance of a native speaker of Japanese. The Japanese text was then transcribed to roman alphabet. The audio samples were split to single sentences and matched to the roman transcriptions using Praat (Boersma & Weenink). Praat scripts were used to identify the approximate location of the vowels, and then final matching was done by hand. For this matching procedure care was taken to use visible changes in wave form or spectrum to identify the beginning and end of segments. Whenever possible the presence of pitch and changes in intensity were used to identify vowels. Here it should be noted that the aforementioned devoicing of high vowels makes this difficult, and in fact devoiced vowels in the presence of a fricative often do not seem to be overtly realized at all. Also a strong tendency towards creaky voice at the end of phrases makes identification of vowels by pitch unreliable. On the other hand the regularity of Japanese timing means that the simple procedure of dividing a phrase by the number of segments often comes close to identifying vowels correctly.

Issues arising in the identification of vowels, aside from the issue of voiceless vowels, were primarily the result of vowel sequences. When sequences of identical vowel arise, either in word + particle combinations (*katsudō-o*) or across words (*nyūsu-o otsutae*), such sequences are generally spoken as one long uninterrupted vocalic segment. Also sequences of [a] with intervening [w] (*kawase*) are sometimes spoken without any phonetic indication of the [w] essentially becoming one long [a] segement. More problematic, as it may affect the quality of the vowel, are sequences such as [ai], either mono-morphemically or across words. While some Japanese phonologists believe such cases should be treated as two vowels, it is often almost impossible to identify any clear boundary, and so they are probably best treated as diphthongs.

Phrases are both semantic and prosodic units. They can be identified on the basis of short pauses made by the speaker as well as a drop in pitch (Pierrehumbert & Beckman, 1988). Overall the number of tokens collected are summarized in (2).

(2)

Sample Nr.	M/F	sentences	phrases	a-final	[a]	final [a]
1 + 2	M	17	67	32	211	32
3	M	3	9	7	50	7
4	F	6	31	12	148	12
Total		26	107	51	409	51

As mentioned, 4 samples were collected spoken by 3 different announcers: two male and one female. The samples consisted of 26 sentences. Sentences consisted of anywhere from 1 to 9 separate phrases. Almost half of all phrases ended in the vowel [a].

Examples of sentence medial phrases endings in [a] are due to various particles:

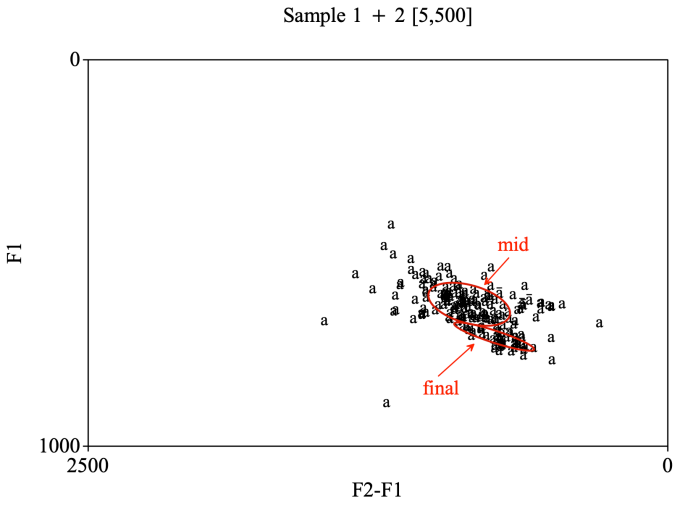
(3) *...yōgisha-wa*
...kēzaisēchōritsu-ga
...hiekondeiru-hoka
...gaishō-no hō-kara

Sentence final instances of [a] in the samples at hand are exclusively forms of verbs ending in */~mashita/*.

After vowels were labelled, Praat was used to perform a spectral analysis. Frames with fewer than 5 formants were eliminated. It was found that for the male speakers a formant ceiling of 5,500 Hz performed best, while for the female ceiling of 7,000 Hz gave the best results.

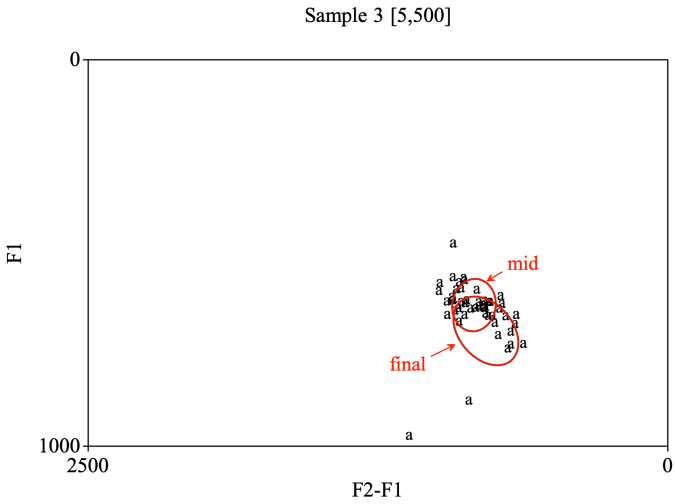
The resulting distribution diagrams are shown below.

(4)



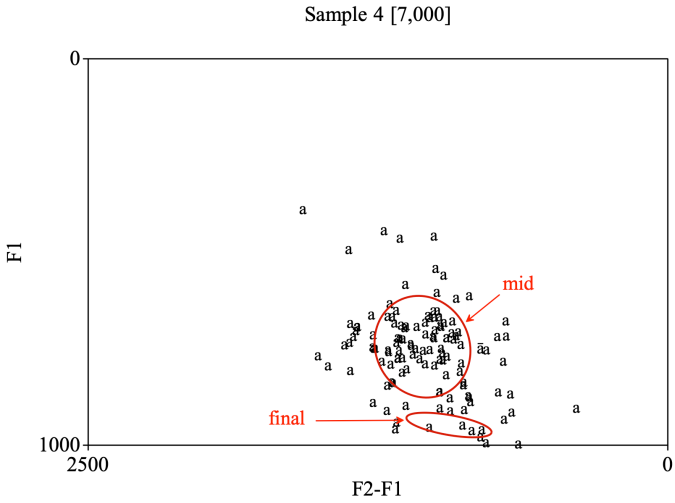
Sample 2 was the largest of the 4 samples, and was analyzed together with Sample 1 which came from the same announcer. As can be seen, the phrase final [a]s cluster tightly at the bottom of the range of the distribution.

(5)



Sample 3 was somewhat short. Here too, the phrase final [a]s were generally clustered near the bottom of the range.

(6)



Sample 4 was the female speaker. In this case, instances of phrase final [a] were consistently found to be considerably lower than the phrase medial [a]s.

While the samples used for this investigation are certainly somewhat limited, the results are quite encouraging, and a more thorough investigation seems warranted.

4. Discussion

The investigation performed in this paper seems to indicate that the phenomenon described is in fact measureable and reproducible. In fact, based on personal observation, NHK newscasters

consistently articulate phrase final [a] in the manner discussed here. In light of these results, follow up should include:

- Gathering of more data from more speakers, including newscasters from other broadcasting companies.
- Statistical evaluation of the data to determine whether the results are significant.
- Evaluation of data from non-newscasters, to determine whether this is indeed specific to newscasters or whether it is found in the speech of others as well.

References

Boersma, Paul and David Weenink. *Praat: doing phonetics by computer.* `https://www.fon.hum.uva.nl/praat/`.

Ladefoged, Peter and Ian Maddieson. 1996. *The Sounds of the World's Languages.* Blackwell.

Pierrehumbert, Janet and Mary Beckman. 1988. *Japanese Tone Structure.* MIT Press.

Vance, Timothy J. 2008. *The Sounds of Japanese.* Cambridge University Press.

(Received: January 10, 2022)