Phonetica
Editor: K. Kohler, Kiel

Separatum
Publisher: S. Karger AG, Basel
Printed in Switzerland

Phonetica *41:* 191–207 (1984)

0000–0000/84/0414–0191$2.75/0

Vowel Variation in Japanese

Patricia A. Keating, Marie K. Huffman

Phonetics Laboratory, Department of Linguistics, University of California, Los Angeles, Calif., USA

Abstract. A language's use of the phonetic vowel space depends not only on how many vowel phonemes the language has, but on how each phoneme varies allophonically across contexts. This study tests the hypothesis that Japanese vowel allophones, measured from a wide range of contexts, will not fill the vowel formant space. This was predicted because Japanese has few vowel phonemes, distributed unevenly in the vowel space, and has no obvious processes of vowel reduction. Formant frequencies of vowel tokens from word lists and from read texts were compared for 7 speakers. These data show that, in Japanese, vowel allophones in prose fill in the vowel formant space more than allophones in word lists do, mainly as a result of centralization of the prose tokens. The use of the total formant space is determined in part by the distribution of phonemes, and in part by allophone centralization.

Introduction

It is well known, especially for English, that vowels vary phonetically across segmental and prosodic contexts. Contextual effects on vowels include nasalization in nasal environments (e.g. 'm*a*n'), lengthening (e.g. 'b*a*d'), shortening (e.g. 'pref*e*rence'), and lip rounding (e.g. 'shr*u*b'). A variety of such effects can co-occur in, for example, a possible pronunciation of 'choose' with [y:] – the vowel is fronted by the surrounding coronal consonants, and lengthened by the following [z]. Thus a listing of the phonetic qualities of the English vowel phonemes gives little hint of the wide range of phonetic variants encompassed across contexts.

Vowel formant variation has been of particular interest. Vowels are more centralized in the vowel formant space in consonant contexts [*Stevens and House,* 1963] and when they are shorter [*Lindblom,* 1963, for Swedish], and they are affected by nonadjacent vowels [*Öhman,* 1966; *Fowler,* 1981]. Other studies of vowel formant variation include *Joos* [1948], *Stevens* et al. [1966], *Tiffany* [1959], *Broad and Fertig* [1970], *Ohde and Sharf* [1975], *Bell-Berti and Harris* [1976], and *Schouten and Pols* [1979, for Dutch]. Most of these have concentrated on

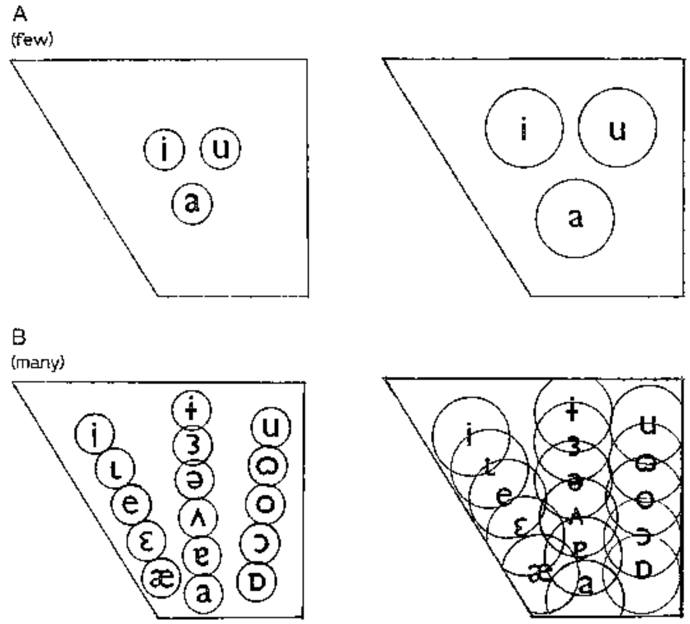

Fig. 1. Comparison of vowel formant spaces in hypothetical languages with few vowels (top row) and many vowels (bottom row), with little variation per vowel (left column) and much variation per vowel (right column).

the second formant, that is, variation due to front-back movement of the tongue and to lip rounding.

Is the situation in English typical of other languages? We note that traditional phonetic descriptions of some languages, such as Russian [*Jones and Ward*, 1969] and Arabic [*Abdel-Massih*, 1981], stress the phonetic variation of vowels, especially in the front-back dimension, while descriptions of other languages do not. If languages do indeed differ in how much their vowels vary, then there are two ways they might do so. One possibility is that all languages use the same total vowel space. In this case the amount of vowel variation would be inversely correlated with the number of vowels in the language, with neighboring vowels filling holes in asymmetrical inventories by varying more. The other possibility is that languages differ in the total amount of vowel space used, due to arbitrary differences in the way their vowels vary.

To visualize the relation between number of vowels, amount of variation per vowel, and their effect on use of the phonetic space, consider figure 1, where two languages with symmetrical vowel systems are illustrated with variation in the two-dimensional $F_1 \times F_2$ vowel space. (Of course, the true vowel space includes other dimensions, as noted above.) One language has few vowels (A, top two diagrams), and the other has many (B, bottom two diagrams). The circle around each vowel symbol represents its range of allophonic variation. The left column shows the phonetic space if vowels each vary little, and the right column shows the space if vowels each vary quite a bit. Language B has so many vowels that regardless of how much each one varies, the space is virtually filled. Language A fills the space only if each vowel varies an extreme amount. Otherwise much of the space remains blank.

Thus we see that if each vowel varies to the point of considerable overlap with other vowels, then in the limit, the number and distribution of vowels in a system becomes irrelevant – any language can fill the entire space. On the other hand, if the amount of variation per segment is somewhat limited, then use of the phonetic space depends on the number of vowels, and how they are distributed, independent of context. *Liljencrants and Lindblom* [1972] and *Disner* [1983] have shown that vowels measured in a single context lie differently in the space across languages, in part correlating with the total number of vowels, but in part idiosyncratically. For example, some languages crowd their phonemes unsymmetrically into some part of the space, leaving other parts rather empty. With more limited amounts of variation per vowel across contexts, these empty spaces should still be visible even when a variety of contexts are taken into account. Thus in sum we see a number of ways in which the phonetic vowel spaces of languages may be distinctive. They may have different numbers of segments (with roughly the same amount of variation), different degrees of variation per segment (with roughly the same number of segments), different distributions of segments (with little variation per segment), or some combination of these factors. However, if the variation per vowel is great enough, any other differences will not matter, and the languages will not differ phonetically in their use of the total possible vowel space.

If any language might be expected *not* to fill up the vowel space, then, it would be a language with few vowels, unevenly distributed, and each showing little phonetic variation. By most accounts Japanese appears to be such a language. Japanese has only five vowels, unevenly skewed towards high front vowels. Japanese also is said to have the interesting property that in certain contexts vowels are deleted altogether, thus eliminating one possible source of qualitative variation. For these reasons Japanese could fail to fill the vowel space with vowel allophones. In this paper we test this hypothesis by studying the vowel formant space of Japanese in detail.

The Vowel System of Japanese

Japanese has five phonemic vowels, each occurring long and short. Sometimes the five vowels are given as /i, e, a, o, u/, but more phonetically precisely they are /i, e, a, o, ɯ/. None are diphthongized. The long vowels can be analyzed as sequences of two short vowels, though such sequences differ in junctural properties from sequences that arise across syllable boundaries.

The acoustic characteristics of Japanese vowels have been studied to some extent. An early study of Japanese vowel formant frequencies is *Umeda* [1957], published in Japanese. She gives spectrographic measurements of four formants for male and female speakers producing CV syllables in sentences and in isolation. In general, the second formant frequencies appear to be slightly higher in the sentences than in isolation, but the differences are not great. The pooled results for F_1 and F_2, plotted like a traditional vowel diagram, look very much like those shown in figure 2, which shows the results of our analysis, to be described below.

A classic reference on Japanese phonetics is *Han* [1962]. In this study formant frequencies are presented for the five vowels 'in various speech environments'. In *Han's* [1962] speech, there is a slight asymmetry in vowel height for front versus back vowels, with the front vowels being somewhat higher than the back vowels, but for other speakers this is not the case. More recently, *Aoki* et al. [1984] give frequencies and bandwidths for F_1, F_2 and F_3 of vowels pooled over a variety of consonant contexts, for 1 speaker. Again, their results look like figure 2. *Aoki* et al. [1984] note that short vowels, which are more common than long vowels, tend not to reach their 'target' values. They identify four contexts that especially favor coarticulatory effects: /a/ between dental consonants, /o/ between any coronal consonants, and any two adjacent vowels, all have their formant frequencies affected; also any vowel followed by a 'syllabic nasal' is nasalized. However, they conclude that much of the variation in their data is unexplained.

As shown by these various studies, and by our own results below, phonetically there are four degrees of F_1 values (i, ɯ, e and o, a) and four degrees of F_2 values (i, e, ɯ and a, o) in these five vowels. The F_1 values correspond relatively straightforwardly to vowel height differences. The F_2 values correspond to the combined effect of differences in vowel backness and vowel rounding. Thus the more centralized position of Japanese /ɯ/ in the formant space than of, for example, English /u/, is surely due in part to the Japanese vowel's being unrounded, but may also indicate some relative tongue fronting as well. Phonologically, only two degrees of height and backness are distinctive.

The most important natural class among the vowels is the high vowels /i/ and /ɯ/, which figure together in several phonological rules [*Haraguchi*, 1984]. One rule that has received some attention is the deletion or devoicing of high vowels under certain conditions. The traditional, but inadequate, statement is that high vowels between voiceless consonants or between a voiceless consonant and pause are 'devoiced' or 'unvoiced' especially when occurring with low tone. *Han* [1962] discusses vowel 'unvoicing' in some detail. She presents spectrograms of her own speech that show shortened and devoiced vowels, and argues that even if not heard by nonnative listeners, the vowels are 'present' in the sense that their allotment of time in the word is taken over by a preceding segment. However, as *Beckman* [1982] shows, these vowels are normally deleted rather than voiceless. *Han* [1962] relates vowel unvoicing to vowel shortness, in that unvoicing is more likely to occur with the inherently shorter high vowels, and at a faster tempo, where any vowel is shorter, and between voiceless consonants, where she claims that again any vowel is shorter. She says that unvoicing is especially common after /s/ and /ʃ/, but uncommon in accented syllables, and its relation to vowel shortening remains somewhat mysterious.

The quasi-systematic deletion of short vowels in Japanese suggests that the usual processes of reduction of shorter vowels will not play a large role in the determination of allophonic vowel quality. Vowels that in other languages would be likely candidates for reduction are expected to delete in Japanese. This process provides another reason to predict relatively little allophonic variation in Japanese vowels. In this study, the hypothesis being tested is that

Japanese vowel formants, measured from tokens occurring in a wide range of contexts, will fail to completely fill the available formant space. Furthermore, the unfilled areas should be relatable to the distribution of the vowels in a neutral context.

Method

Speakers

7 young male speakers of Japanese from Tokyo were recorded. A Japanese graduate student in linguistics who also served as a speaker ascertained that each speaker was in fact a speaker of a Tokyo dialect. All speakers were paid volunteers.

Materials

Two types of materials were selected for this study. Three sets of words each contrasted the five short vowels in different contexts, as shown in table I: (1) high-toned between /h/ and /b/; (2) low-toned between /b/ and pause; (3) high-toned between /b/ and pause. In addition, five prose passages were taken from Japanese textbooks and magazines. The amount of text chosen was determined by our goal of recording 100 tokens of each short vowel by each speaker, taking into account where vowels were likely to be deleted in normal fluent reading. The purpose in having so many tokens was to compensate for the lack of control over the texts' phonetic characteristics, especially idiosyncratic aspects of reading prosody. In general, the vowels occur in reasonably well-distributed consonant environments. However, for two vowels, some unevenness is seen, in that /o/ and /e/ are found mainly after apicals; furthermore, /a/ is found in a disproportionately small number of morphemes, rather than in a wide range of items.

Recordings

Audio recordings were made of the 7 speakers reading these materials in a sound-treated booth in the UCLA Phonetics Lab on an Ampex recorder. All materials were presented in Japanese orthography; a transliteration was also provided to the experimenters by a native Japanese. The word lists were read first, in a pseudorandomized order. Since there were 15 words read twice, a total of 30 word list items were produced by each of the 7 speakers. Then the texts were read, at a self-selected rate and loudness.

Table I. Word list tokens for the five short Japanese vowels

	Tones		
	HL	HL	LHH
i	h*i*bi	tab*i*	wasab*i*
e	h*e*bi	nab*e*	otab*e*
a	h*a*bu	kab*a*	anab*a*
o	h*o*bo	yab*o*	otsub*o*
ɯ	h*u*bo	kob*u*	manab*u*

Analysis

In analyzing the data, we began by digitizing approximately 2 s of speech at a 10 K sampling rate in the UCLA Phonetics Lab's PDP 11/23 computer. The Waves speech analysis system was used to view the waveforms. The experimenters used the transcription of the texts to locate individual vowel tokens. For each candidate vowel, computer listening and viewing facilities were used to select those tokens long enough for LPC analysis, i.e. to determine whether there was at least 30 ms of voiced oral vowel with relatively steady amplitude and constant auditory quality. If there was, that vowel chunk was stored for later analysis. If there was not, the vowel was eliminated. Thus there are three categories of vowels eliminated at this stage: vowels not present either auditorily or acoustically; vowels that were present but too short to analyze, such vowels also sometimes being inaudible (though on a spectrogram, they would be quite visible and usually measurable); vowels present but not usable for some other reason, such as nasalization or noise. In cases of fairly long vowels, most of the steady state was generally used.

The stored vowel tokens were then analyzed using LPC to find formant frequencies. A 30-ms Hamming window was moved over the sample in 5-ms steps. At each step 14 LPC coefficients were found by the autocorrelation method, and roots of a polynominal defined by these coefficients were used to compute formant frequencies and bandwidths. The lowest five candidate formant frequencies and bandwidths were printed out for each step, and from these, the experimenter decided where to choose frequency values for F_1, F_2 and F_3. First, inspection of bandwidths was used to eliminate spurious formants. Then a single mean frequency for each formant was obtained by averaging the two or three steps, for even and odd numbers of steps, respectively, in the middle of the steadiest portion of the token. Any tokens for which measurements could not be determined with confidence were eliminated. The formant frequency data, then, consist of a single value for each of F_1, F_2, and F_3 for each token, those values being means over a 31- to 36-ms steady-state portion. Each speaker's values were compiled together with their source words for further analysis. For each speaker, the same number of tokens of each vowel was used, determined by the vowel having the smallest number of successfully measured tokens (typically /ɯ/). Recall that our goal was 100 tokens per vowel, but some speakers deleted /i/ and /ɯ/ so freely that we could rarely get more than 90 tokens of each of these vowels, and therefore for all the vowels. Thus the number of tokens differs across speakers, ranging from 59 to 100. It is also the case that, because of individual differences in deletion, the particular tokens (i.e. words in the text) that ended up being analyzed differed slightly across speakers.

Spectrograms were also made of some of the read texts, for several reasons. They helped in identifying vowel formants, for example, in locating steady states in vowel sequences, and in interpreting LPC outputs. They were used for distinguishing vowel devoicing from deletion, and for measuring vowel duration.

Results and Discussion

Effects of Excluded Vowels

One of the reasons we expect Japanese vowels not to vary much acoustically in prose is that there is no reported vowel reduction. Instead, certain vowel tokens are expected to devoice, delete, or shorten so that formant measurements will not be available. Indeed, many vowels in our texts were eliminated in those ways. The literature suggests that particular consonant environments are associated with these processes, which would mean that particular kinds of coarticulatory effects on vowels would be systematically absent. However, in general that was not the case for our corpus. First, true voiceless vowels occurred only at the ends of utterances, when one or more entire syllables were devoiced, and this process affected all vowels. Second, deletion and shortening appear to be two degrees of the same phenomenon, so that overall, no vowels were uniformly absent from particular contexts. Rather, in many repetitions of any given word, a vowel will sometimes be deleted, sometimes be extremely short, and sometimes be neither. Furthermore, our speakers differ in their liberality in this regard. That is, deletion is just unreliable enough that most sequences do show up at least sometimes. At the same time, high vowels are so much shortened even when not deleted that a majority of high vowels in our sample do come from voiced environments. As for other vowels, fricative environments may be somewhat underrepresented: voiceless fricatives favor shortening, and voiceless fricatives are much more common than voiced in our texts.

At least two other classes of vowels were excluded from the analysis to be presented, and we should consider briefly how they could have affected our results. First, some vowels that were present in the speech signal were too short for the LPC routine to analyze. To determine whether such very short vowels differ from the included vowel tokens, spectrograms were used to measure formant frequencies of a subset of such

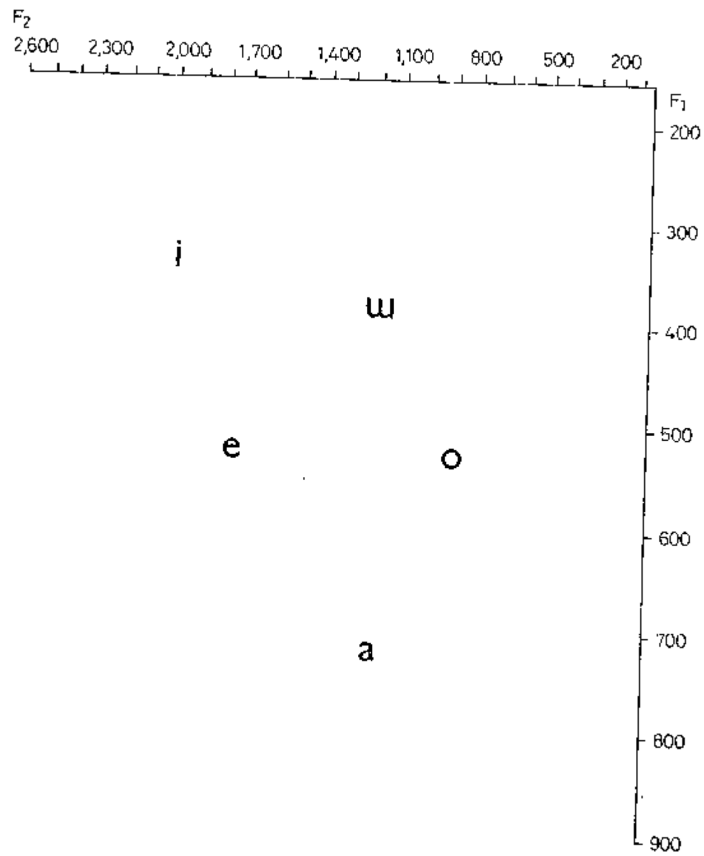

Fig. 2. Mean values for each vowel in word lists, all speakers. All values are in Hz.

very short vowels. Those tokens were found to lie inside the speaker's vowel spaces for the longer tokens. That is, accepting the limits of the LPC routine for formant measurement appears not to have misrepresented the range of vowel variants.

Second, phonemically long vowels were not included. There were far fewer such vowels than short ones in the texts. Furthermore, as it happens, most of the long vowel tokens in our texts were /oo/, with some /ɯɯ/, and otherwise very few tokens of other vowels. Almost all of the long /ɯɯ/ tokens follow palatal consonants. Thus the inclusion of long vowels would have been uneven across the five vowel phonemes. A study of the formant frequencies of a subset of long vowels for 3 of the speakers indicates that for some speakers, inclusion of long /oo/ and /ɯɯ/ would have affected the vowel spaces somewhat. We will return to this point in the discussion below.

Vowels in Word Lists

When the 30 word list tokens for each speaker are compared, a quite similar pattern is seen across speakers. Figure 2 shows the mean values for the five vowels in the $F_1 \times F_2$ space for all speakers. Figure 3 shows the mean values and the individual tokens underlying the mean for one of the speakers. The axes of the graphs have been

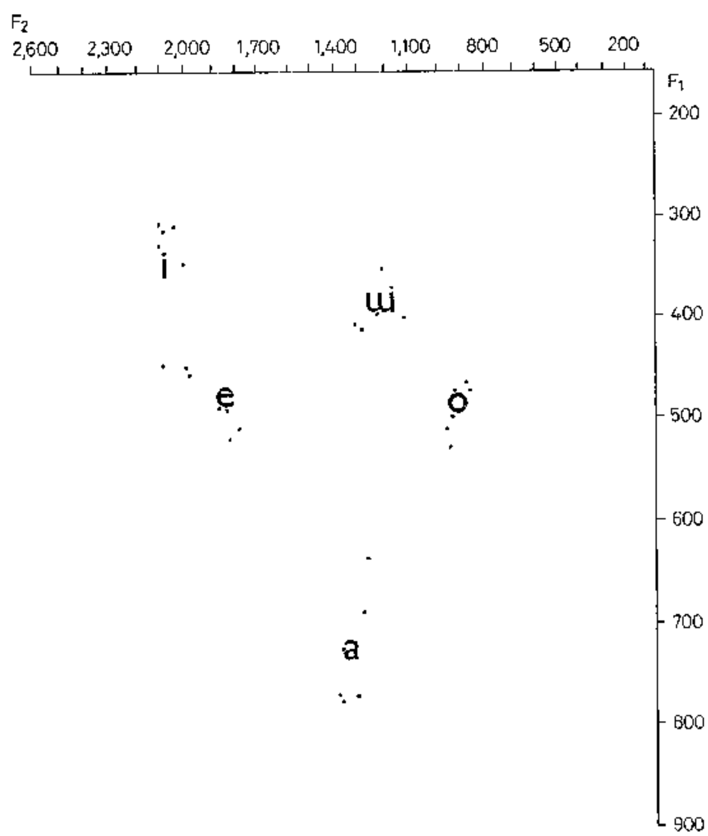

Fig. 3. Individual tokens and mean value for each vowel in word list, speaker 1.

rotated so that the graph corresponds to a traditional chart of tongue position.

The mean vowels /e/, /o/, /ɯ/, and /a/ form a diamond pattern, with /e/ and /o/ equal in height, and /ɯ/ and /a/ equal in backness. /i/ is both higher and fronter than any other vowel, giving the set of five vowels a skewed distribution towards the high front part of the space. The amount of variation shown by individual tokens of each vowel is not sufficient to fill in the vowel space; certain areas of the space are empty. These include the regions for high back round [u] or [o], mid central [ʌ] or [ə], low front [æ], and low back [ɑ]. Comparison of the 7 individual speakers shows that they divide into two groups. 5 speakers have the two mid vowels closer to the two high vowels and farther from the low vowel, while 2 speakers have the two mid vowels closer to the low vowel. However, this difference is not great, and all speakers use a similar total range of formant space for the set of vowels.

On the basis of this distribution of vowels, it can be seen that some vowels are already rather close to other vowels (e.g. /i/ and /e/), while some vowels are surrounded by empty regions of the vowel space (e.g. /a/ and /ɯ/). Given this, we would predict that, if vowels vary so as to fill the available space, /i/ and /e/ should

Table II. Means and standard deviations in Hertz of three formants for vowels in prose

Speaker	Number of tokens per vowel	Vowel	F_1 mean	F_1 SD	F_2 mean	F_2 SD	F_3 mean	F_3 SD
1	81	i	347	39	2,001	98	2,762	203
		e	467	45	1,808	121	2,523	127
		a	614	80	1,517	147	2,401	190
		o	458	47	1,196	205	2,352	165
		ɯ	386	43	1,493	242	2,307	160
2	93	i	348	45	1,753	147	2,749	269
		e	473	40	1,586	99	2,335	155
		a	605	61	1,309	99	2,191	210
		o	472	57	1,144	150	2,070	198
		ɯ	398	47	1,352	206	2,063	265
3	100	i	350	47	1,843	133	2,637	154
		e	502	49	1,634	133	2,432	71
		a	665	43	1,208	95	2,382	163
		o	503	37	1,043	152	2,347	232
		ɯ	403	45	1,345	203	2,248	178
4	83	i	364	53	2,120	180	2,971	200
		e	453	70	1,905	110	2,760	101
		a	656	94	1,596	112	2,562	132
		o	486	48	1,228	183	2,605	164
		ɯ	410	51	1,482	281	2,541	168
5	74	i	379	42	1,973	173	2,686	169
		e	461	38	1,692	162	2,527	83
		a	572	54	1,368	135	2,447	121
		o	484	55	1,227	181	2,482	130
		ɯ	423	43	1,504	207	2,373	159
6	59	i	357	47	2,140	139	2,979	300
		e	488	49	1,773	146	2,607	106
		a	656	76	1,395	115	2,397	162
		o	473	46	1,072	111	2,494	139
		ɯ	402	63	1,402	290	2,388	122
7	96	i	371	38	1,851	101	2,572	183
		e	481	40	1,644	124	2,399	85
		a	647	52	1,290	97	2,343	127
		o	488	32	1,044	106	2,445	148
		ɯ	413	40	1,357	207	2,248	137

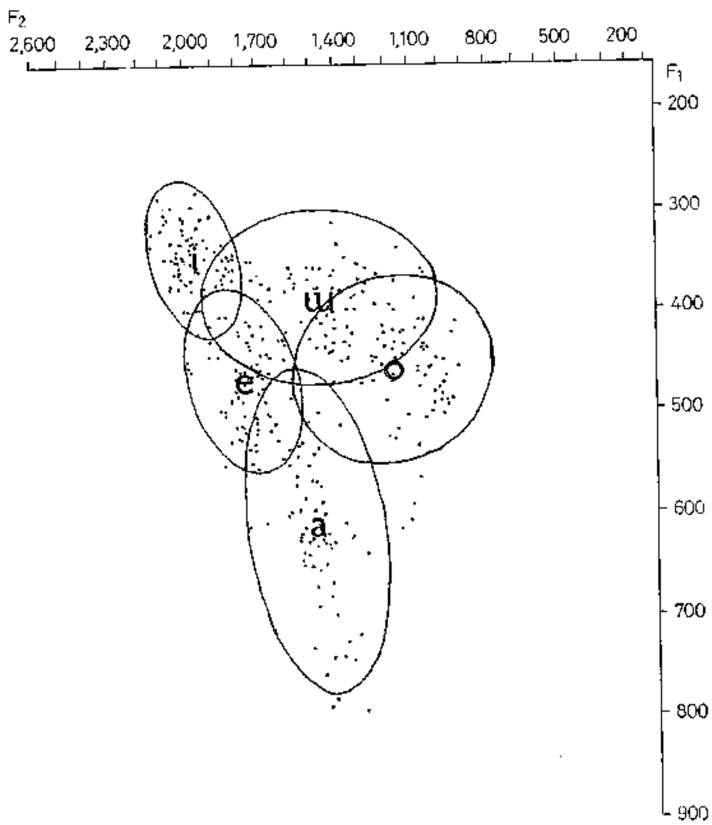

Fig. 4. Individual tokens and mean value for each vowel in prose, speaker 1, same speaker as in fig. 3.

vary the least, /a/ the most, and /ɯ/ more than /o/. Let us see, then, how the vowel tokens in prose actually do vary.

Vowels in Prose

Table II summarizes the mean formant frequency data for each prose vowel for each of the 7 speakers. Figures 4 through 10 show the means and individual tokens for each speaker. (Figure 4 represents the speaker whose word list tokens are shown in figure 3.) Ellipses are drawn around the tokens, with radii of two standard deviations along the two principal components. As can be seen, speakers differ in the amount of variation per vowel and in the size of the resulting formant space. Thus, for example, the speaker represented in figure 6 has relatively less variation for each vowel; the ellipses are small, they overlap little, and the center of the formant space is empty. On the other hand, the speaker represented in figure 7 has relatively more variation per vowel; the ellipses are large, they overlap more, leaving no gaps in the center, and the total formant space is larger than other speaker's.

Is the vowel formant space filled in by variation of vowels in prose? It would seem that the center of the space is somewhat filled by allophonic variation , while the periphery is not. The vowel /ɯ/ spreads in-

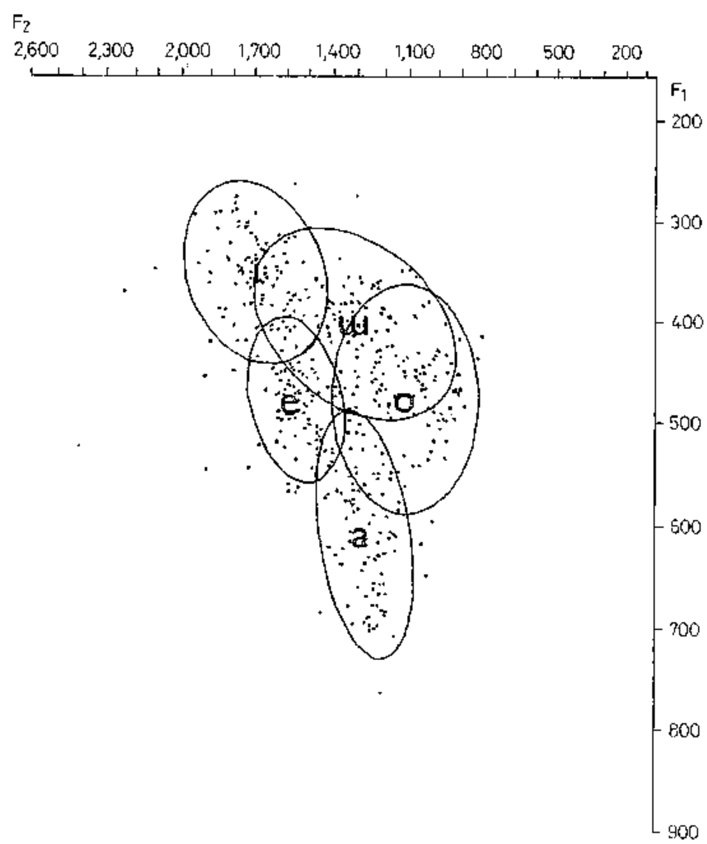

Fig. 5. Individual tokens and mean value for each vowel in prose, speaker 2.

ward in the space because its major variation is in F_2, while /a/ spreads upward because its major variation is in F_1. Thus /ɯ/ provides higher central allophones, while /a/ provides lower central allophones. At the same time, the other areas of the space that were empty in the word lists remain basically empty in the prose, namely, the corners other than [i]. Some spreading of /ɯ/ and /o/ towards [u] is seen, as is some spreading of /e/ and /o/ towards [ɛ] and [ɔ], respectively, but the extreme corners remain clearly empty. The distribution of allophones preserves the skewing of the vowel phonemes towards the high front region of the space.

These results may be at odds with those of *Dorsey and Bernstein* [1981], who compared the overall phonetic spaces for all segments of Japanese and American English. 2 Japanese talkers read 4 min of text, and 2 Americans read sentences. 12 LCP coefficients were computed using a 30-ms window stepped along at 20-ms intervals for the entire speech corpus (not just the vowels). The two languages' LPC coefficient spaces were compared to see whether Japanese, with fewer phonemes, used less space than English, with more phonemes. *Dorsey and Bernstein* [1981] tentatively concluded that the two languages use the *same* size LPC coefficient space, despite the dif-

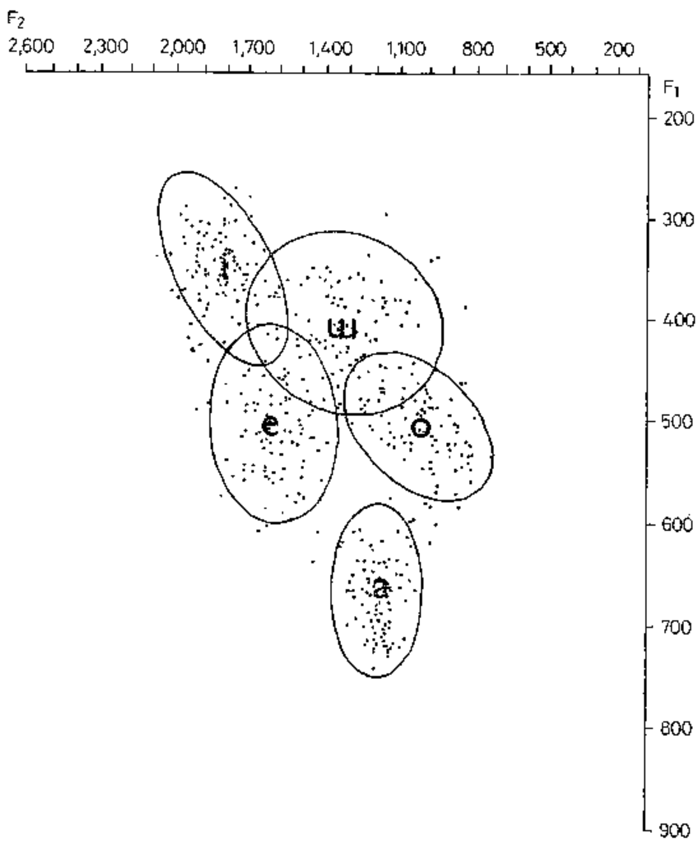

Fig. 6. Individual tokens and mean value for each vowel in prose, speaker 3.

ference in number of phonemes. Since our results for Japanese vowels show empty areas of the formant space that are surely not empty in English, we would not want to conclude that the two languages use similar vowel spaces. However, *Dorsey and Bernstein's* [1981] claims are not phrased in terms of the formant space, so cannot be directly evaluated against our data. In addition, there are a few factors in *Dorsey and Bernstein's* [1981] design that could have influenced their results. Their study was based on all segments, not just vowels; it used only 2 speakers of each language, and the type of material differed (prose from a novel for Japanese, sentences for English), such that the Japanese might have used a greater variety of prosodic contexts.

It is perhaps no surprise that vowel variation found in our Japanese data involves mainly centralization, especially in view of the vowel shortening in our corpus and the known relation of vowel duration and centralization [*Lindblom*, 1963]. The small amount of data on Russian vowel variation presented in *Keating* et al. [1984] indicates that Russian, another five-vowel language, also does not exhibit much peripheral spreading. Our Japanese data illustrate what happens to the phonetic vowel space when this trend is the major determinant of vowel variation: while new vowel qualities

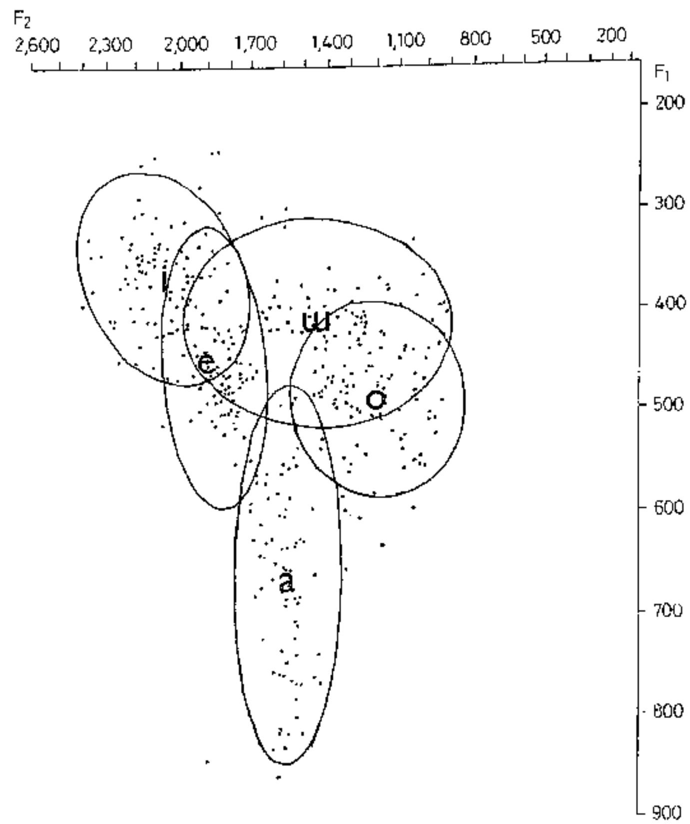

Fig. 7. Individual tokens and mean value for each vowel in prose, speaker 4.

may arise through allophonic variation, the overall size and shape of the formant space is constrained by the phonemic inventory. Peripheral vowels lacking from that inventory do not seem to arise through allophonic variation. Of course, evidence from Japanese does not mean that a language like, e.g., Tausug [*Disner*, 1983], with rather centralized high vowels, could not in principle display higher allophones through phonetic variation. It simply means that acoustic evidence for such phonetic variation has not been found.

At the same time, we must note that we have excluded cases of long vowels, which might be expected to provide somewhat more peripheral variants than the short vowels. As mentioned above, in our corpus the long vowels consist almost exclusively of /oo/ and /ɯɯ/, the latter in few contexts. For 3 of the 7 speakers we considered how inclusion of long vowels could have affected the appearance of the overall formant space. For one speaker, no difference could be seen with either long vowel. For the second speaker, the /ɯɯ/ tokens showed no difference, but the /oo/ tokens were more peripheral in the space, being both slightly higher and backer. For the third speaker, the /oo/ tokens behaved similarly, and in addition the /ɯɯ/ tokens showed a difference. However, that differ-

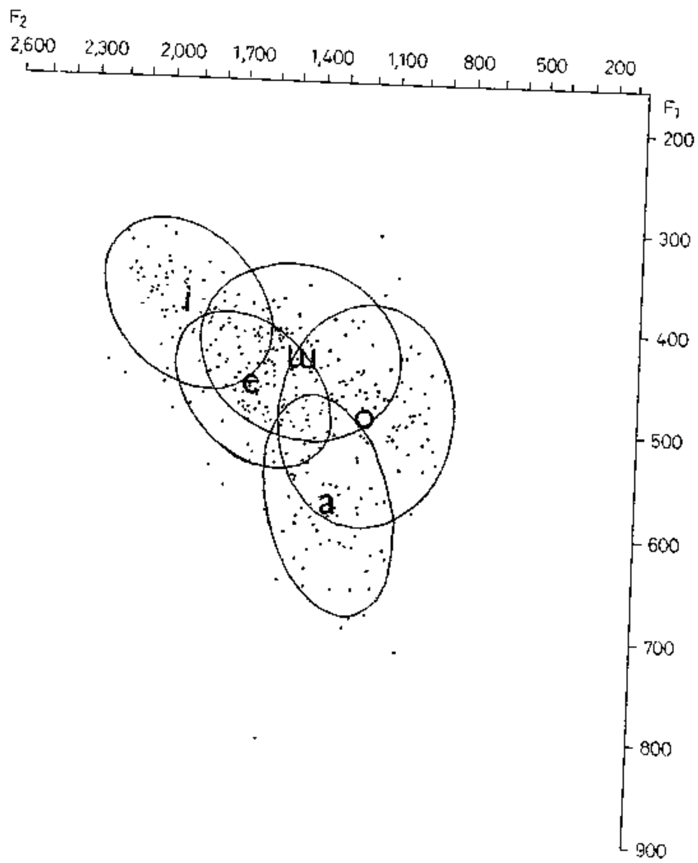

Fig. 8. Individual tokens and mean value for each vowel in prose, speaker 5.

ence did not affect the periphery of the formant space: the /ɯɯ/ tokens, mainly following palatal consonants, had high F_2 values, and therefore filled in more of the high central part of the formant space. Thus, in sum, given the small number of long vowels and their unequal frequencies in the corpus, they were best excluded from the analysis, but their inclusion would have resulted in a slightly higher and backer periphery of the space for some speakers, due to differences between /o/ and /oo/.

Do the vowels nearer the empty regions in the space vary more than the vowels near other vowels? To some extent this is the case. /ɯ/, with the high central, high back, and mid central regions nearby, does vary the most. It covers a larger area than other vowels for each speaker, and its position in the space differs more across speakers. However, that variation is mainly in one direction. For the speakers with word list /a/ relatively far from the mid vowels, /a/ also varies a great deal, but again mainly in one direction. For other speakers, /o/ shows more variation, and indeed /o/ is adjacent to some empty regions – but the variation does not extend much into the empty high back region; rather it contributes to filling in the center region.

In sum, while it may be true that vowels near phonologically empty regions show

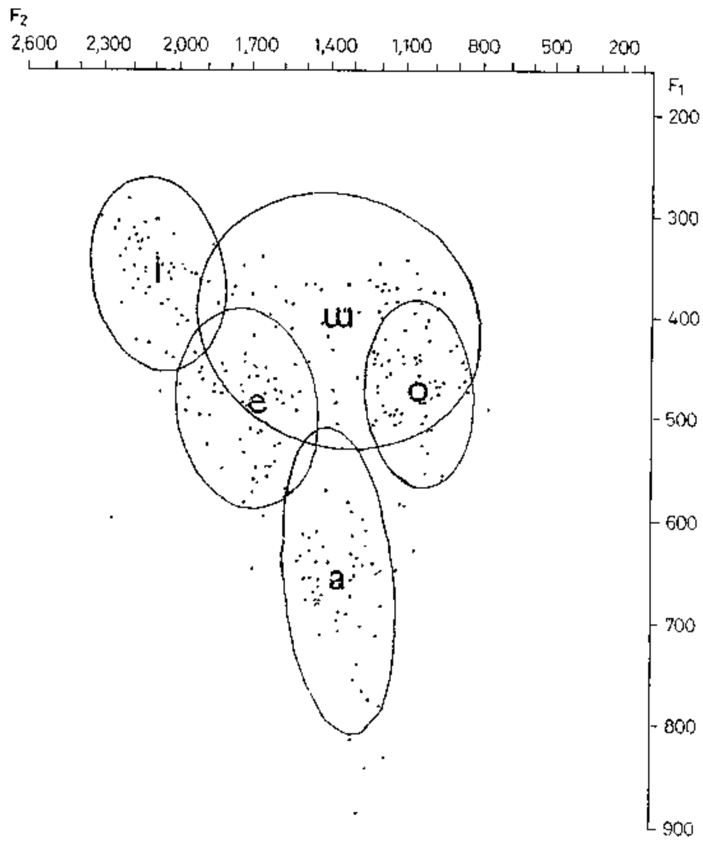

Fig. 9. Individual tokens and mean value for each vowel in prose, speaker 6.

more variation, by and large that variation does not seem to be directed at filling those empty regions. Thus it is premature to conclude that the amount of variation exhibited per vowel is simply a function of its position relative to the other vowels of the system.

Conclusions

Data on the distribution of tokens of Japanese vowels in word lists versus uncontrolled prose have been presented. Those data show that allophonic variation produces a vowel formant space that is more filled in for prose than for word lists. This variation consists mainly of centralization, with the center of the space largely filled in for most speakers. The more peripheral regions of the space are less affected by the allophonic variation. Thus the use of the total space is determined in part by the distribution of vowels in a neutral context, and in part by allophonic centralization.

Acknowledgements

This work was supported by grant BNS 80-23110 from the National Science Foundation to *Peter Ladefoged*. We thank him and *Mary Beckman* for their comments on the manuscript; *Noriko Umeda*

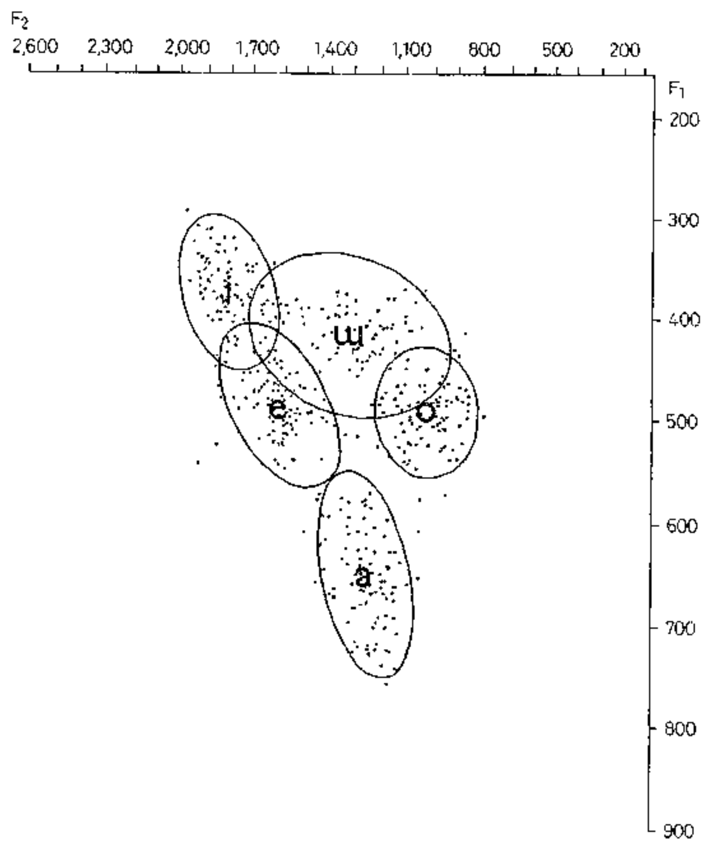

Fig. 10. Individual tokens and mean value for each vowel in prose, speaker 7.

for making available an annotated version of her article in Japanese; and *Shoichi Iwasaki, Ellen Jackson, Jack Martin, Mari Sakaguchi* and *Yuria Wagatsuma* for help with conducting the experiment.

References

Abdel-Massih, C. T.: An introduction to Egyptian Arabic (University of Michigan Press, Ann Arbor 1981).

Aoki, C.; Klatt, D. H.; Kawasaki, H.: Analysis and Synthesis of Japanese Speech. 107th Meet. Acoust. Soc. Am., Norfolk 1984.

Beckman, M.: Segment duration and the 'mora' in Japanese. Phonetica *39:* 113–135 (1982).

Bell-Berti, F.; Harris, K. S.: Some aspects of coarticulation. Haskins Lab. Status Rep. Speech Res., SR 45/46, pp. 197–204 (Haskins Laboratories, New Haven 1976).

Broad, D. J.; Fertig, R. H.: Formant-frequency trajectories in selected CVC-syllable nuclei. J. Acoust. Soc. Am. *47:* 1572–1582 (1970).

Disner, S.: Vowel quality: the relation between universal an language specific factors. UCLA Working Papers in Phonetics, No. 58 (1983).

Dorsey, E.; Bernstein, J.: Inter-speaker comparison of LPC acoustic space using a minimax distortion measure. Proc. 1981 ICASSP, pp. 16–19 (IEEE, New York 1981).

Fowler, C. A.: Production and perception of coarticulation among stressed and unstressed vowels. J. Speech Hear. Res. *46:* 127–139 (1981).

Han, M. S.: Japanese phonology: an analysis based upon sound spectrograms (Kenkyusha Ltd., Tokyo 1962).

Haraguchi, S.: Some tonal and segmental effects of vowel height in Japanese; in Aranoff, Oehrle, Language sound structure, pp. 145–156 (MIT Press, Cambridge 1984).

Jones, D.; Ward, D.: The phonetics of Russian (Cambridge University Press, Cambridge 1969).

Joos, M.: Acoustic phonetics. Language *24:* 1–136 (1948).

Keating, P.; Huffman, M.; Jackson, E.: Vowel allophones and the vowel-formant phonetic space. UCLA Working Papers in Phonetics, No. 59 (1984).

Liljencrants, J.; Lindblom, B.: Numerical simulation of vowel quality systems: the role of perceptual contrast. Language *48:* 839–862 (1972).

Lindblom. B.: Spectrographic study of vowel reduction. J. Acoust. Soc. Am. *35:* 1773–1781 (1963).

Ohde, R. N.; Sharf, D. J.: Coarticulatory effects of voiced stops on the reduction of acoustic vowel targets. J. Acoust. Soc. Am. *58:* 923–927 (1975).

Öhman, S.: Coarticulation in VCV utterances: spectrographic measurements. J. Acoust. Soc. Am. *39:* 151–168 (1966).

Schouten, M. E. H.; Pols. L. C. W.: Vowel segments in consonantal contexts: a spectral study of coarticulation. J. Phonet. *7:* 1–23, 205–224 (1979).

Stevens, K. N.; House, A. S.: Perturbation of vowel articulations by consonantal context: an acoustical study. J. Speech Hear. Res. *6:* 111–128 (1963).

Stevens, K. N.; House, S.; Paul, A. P.: Acoustical description of syllable nuclei: an interpretation in terms of a dynamic model of articulation. J. Acoust. Soc. Am. *40:* 123–132 (1966).

Tiffany, N. R.: Nonrandom sources of variation in vowel quality. J. Speech Hear. Res. *2:* 305–317 (1959).

Umeda, N.: Spectrographic study of Japanese vowels. Technical Report No. 579, Electrical Communication Laboratory (Nippon Telegraph and Telephone, Inc., 1957).

Received: October 17, 1984
Accepted: November 14, 1984

Patricia A. Keating,
Department of Linguistics,
UCLA,
Los Angeles, CA 90024 (USA)