PHONEME FIGURES OF SUSTAINED ORAL VOWELS BY TWO-DIMENSIONAL REPRESENTATION

(II) PHONEME POSITION OF VOWEL-PENTAGON

TERUO FUKUMURA and KENZO HARA

Department of Electrical Engineering

(Received October 30, 1954)

In the Report I, we tried to calculate the two formant positions for five vowels of five subjects, and to plot them on the plane of two-dimension. We reported there on their deviations and their movement with pitch. Here, we will envelope these scattered points by the boundary lines, distinguishing each vowel from others. Finally we characterize them by summarizing them into proper divisions of plane according to the kind of vowel and the variety of subject. We also attempt to propose a pentagon representation for Japanese vowels.

Representation of Difference Boundaries

By enveloping the boundary points for every group of vowels, which are scattered notwithstanding the subject, we obtain the figure in Fig. 5. The facts found in this figure, are that the five groups of vowel-sound are localized in proper domains with considerable clearness, except the region of the border between "A" and "O." This means that Japanese vowel-sounds, uttered according to five phoneme values, can be successufully specified to such a extent by this two-dimensional representation method. To distinguish them quantitatively, several linear lines, parallel to the line of $f_2-f_1=0$, are drawn, and they fairly run through the gap areas outside the closed boundaries that belong to concerned vowels. These are shown also in Fig. 5. From this result it is seen that the five vowels fall into the following band areas of f_1-f_2 plane respectively, that is;

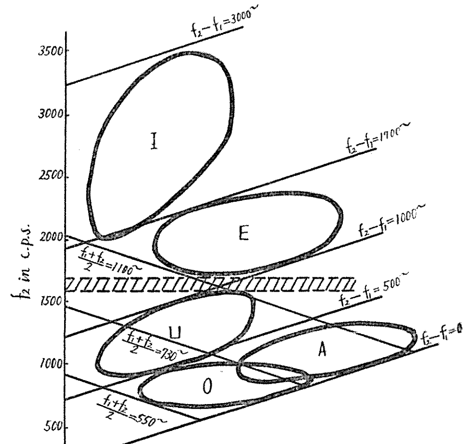


FIG. 5. Vowels bounded by difference lines.

$$f_2-f_1=1{,}700\sim\ <\mathrm{I}<f_2-f_1=3{,}000\sim,$$
$$f_2-f_1=1{,}000\sim\ <\mathrm{E}<f_2-f_1=1{,}700\sim,$$
$$f_2-f_1=500\sim\ <\mathrm{U}<f_2-f_1=1{,}000\sim,$$
$$f_2-f_1=0\sim\ <\mathrm{O},\ \mathrm{A}<f_2-f_1=500\sim.$$

We can draw another sort of boundary line than those explained above. That is:

the following three lines; the boundary line between "O" and "A," the lowest limit of five vowels that contacts with "U" and "O," and one more auxiliary line that demarcates nearly the group "I," "E" and the group "U," "O," "A" are drawn. These are respectively expressed by

$$\frac{f_1+f_2}{2}=730\sim,$$

$$\frac{f_1+f_2}{2}=550\sim,$$

$$\frac{f_1+f_2}{2}=1{,}100\sim.$$

By the way as for the boundary, we can add one more point. Between "E" and "U," at the position of about 1650~ of f_2, the area, where any vowel never appear, is found, and the narrow band through this area, which is shown in Fig. 5 by hatching, separates the vowels into two parts, namely, the group "I," "E" and the group "A," "O," "U." The position of this band corresponds to the position of the so-called vocal glen.[1] In the border regions, especially between 'A" and "O," there are considerable overlapping parts. To check this, an example of articulation testing-score of five sustained vowels, uttered at the pitch of 280~ for five subjects, is presented in Table 2. Referring to Fig. 2 of Report I, it is seen that any vowel, that happens to fall in these border parts tend to have poor clearness, though the deteriorated score of vowel "A" of Sub. H.O. is not reasonably explained. It is infered then that, by correcting the articulatory condition of the subject, the boundary may become still clearer.

TABLE 2. Articulation Testing-Scores of Five Vowels for Five Voices.
(The sustained vowels are uttered at the pitch of 280~.)

Voice→ Vowel↓	H.O. (%)	Y.H. (%)	H.H. (%)	S.M. (%)	T.F. (%)	Mean (%)
"I"	100	100	100	100	100	100
"E"	100	100	100	100	79	95.8
"A"	70	100	100	78	34	76.4
"O"	93	88	76	94	75	85.2
"U"	100	100	100	83	100	96.6

Representation of Interval Boundaries

Instead of the method of linear boundary of f_2-f_1=constant, another method of interval boundary f_1/f_2=constant is proposed, and this is shown in Fig. 6, where the boundary lines which satisfy this mathematical condition, are drawn to differentiate each vowel from others. Then the five vowels fall into their own intervals as follows;

$$f_1/f_2=1/4<\mathrm{I}<f_1/f_2=1/10,$$

$$f_1/f_2=1/2<\mathrm{E,\ U}<f_1/f_2=1/4,$$

$$f_1/f_2=1/1<\mathrm{A,\ O}<f_1/f_2=1/2.$$

Abiding by the criterion of f_1/f_2 interval, the vowels "E," "U" and "A," "O" are arranged in two domains respectively. Two vowels in one of such domains differ from each other only in respect to the absolute values of f_1 and f_2, not to their relative ratio. These can be tested subjectively by means of the rotational synchronous distortion (RSD)* due to out-of-synchronism between recording- and reproducing systems; that is to say, if the rotational ratio as to the reproduction- and recording-revolutional number is larger than 1, possible mishearing from "O" to "A" and from "U" to "E," and *vice versa*, are assumed. This is quite verified experimentally as illustrated in Fig. 7,† where the abscissa represents the rotational ratio as an index of missynchronism, and the ordinate shows percentages of mishearing that takes place. It is very significant and instructive to see that such reciprocal mishearings, which are fairly tested by the distortion RSD, can come forth only in these two cases among all conceivable combinations of five Japanese vowels.

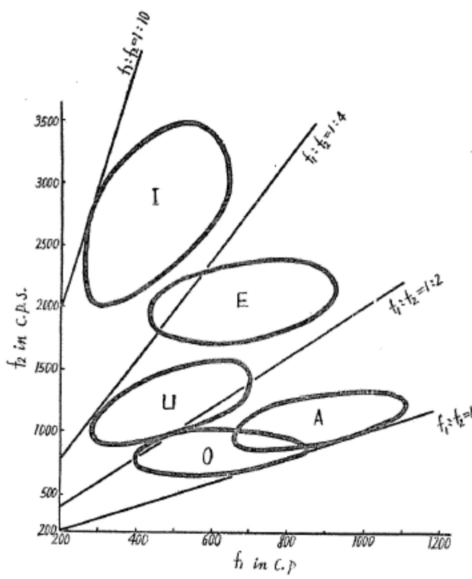


FIG. 6. Vowels bounded by interval lines.

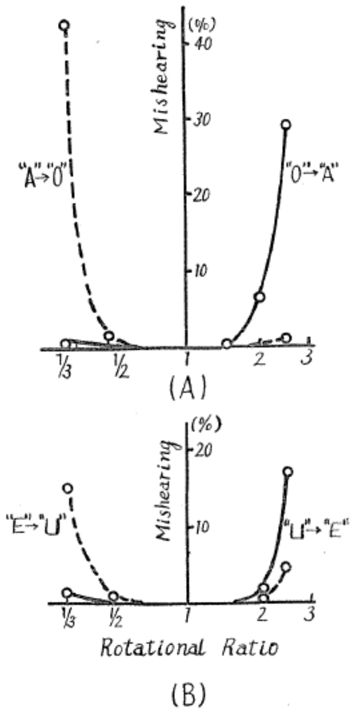


FIG. 7. Mishearing within vowel group by rotational missynchronism, (*A*) for group of "A," "O," (*B*) for group of "E," "U."

Polygonal Representation of Vowels

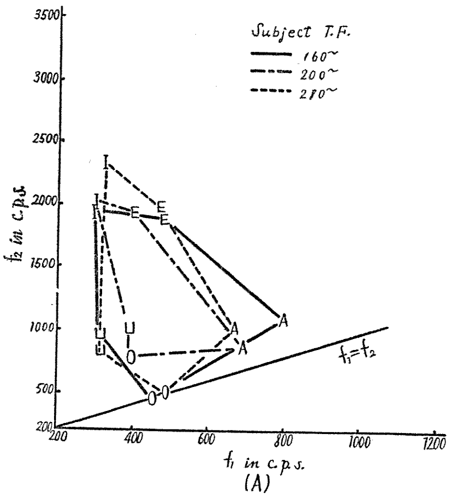
To see more closely the relative position of phoneme figures, the positions of each vowel are successively connected by a closed curve in the order of I-E-A-O-U. And it is seen that the curve tends to form a sort of polygon. As for Japanese

* RSD is the distortion of band transition with constant interval. CSD is the same but constant difference.

† About the quality study of rotational synchronous distortion, an elaborate work has been thoroughly done in our laboratory, and the data are unpublished.

vowel, we would call it "vowel-pentagon."* These are shown in Fig. 8(A) for Sub. T.F., Fig. 8(B) for Sub. S.M. and Fig. 8(C) for Sub. H.H., where the transitions of vowel pentagon with pitch are indicated characteristically at the same time. Those which are shown in Fig. 8, are considered as the most typical cases; where, according to the pitch rise, the pentagon shows little movement in the case of T.F., tending to converge in its area in the case of S.M., and moving its position in the higher direction of the f_1-f_2 plane in the case of H.H.

The effects of subject upon the relative position of the pentagon representation are also shown in Fig. 9, where, for each subject, two typical pitches in natural conversation are used; that is, the pentagons of pitch of 160~ and 200~ for males, and 280~ and 420~ both for females and for child are averaged there. In this case, the axes are both scaled logarithmically. It is interesting to notice that the pentagons are arranged generally in the order of stature of the subject. Disorder is mainly caused by the wide area of H.H.'s. Speaking about vowels, a confusion of arrangement



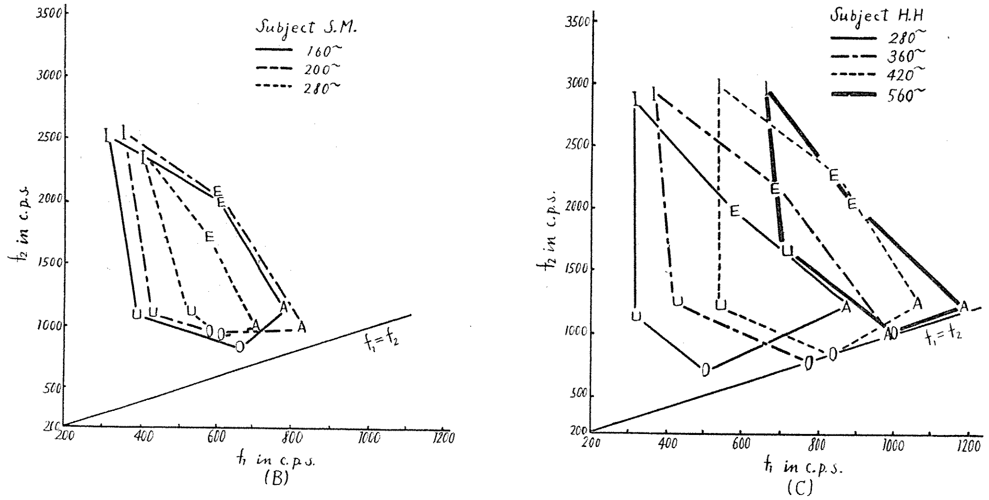


FIG. 8. Vowel-pentagons of different pitches;
(A) for Sub. T.F., (B) for Sub. S.M., (C) for Sub. H.H.

* The reason why the phonemic configuration of Japanese vowel is represented by vowel-pentagon instead of vowel-square as in other languages will be explained mainly by the particularity of Japanese vowel "U" which is uttered generally with a mouth pursed up, and which is specially differentiated from the European vowel "U" uttered with a rounded mouth.

mostly happens in the case of vowels "O" and "A." And it is to be noticed that these have the smallest value of $f_2 - f_1$ among the five vowels.

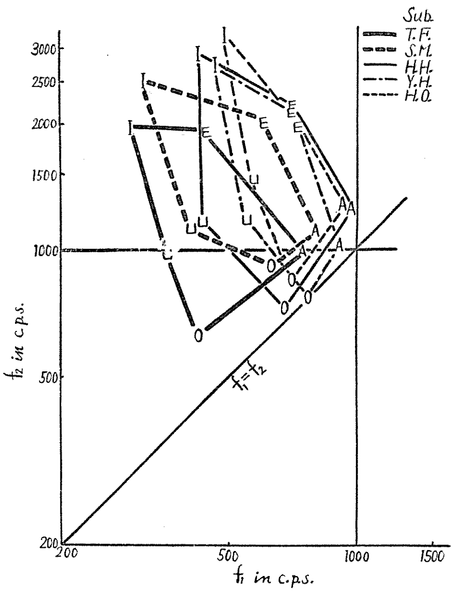


FIG. 9. Relative position of vowel pentagon between subjects, where pentagons are averaged with two pitches which seem to be representative in natural conversation ; 160~ and 200~ for male Subs. T.F. and S.M., 280~ and 420~ for female Subs. H.H. and Y.H. and also for a child H.O.

Comparison of Japanese Vowels with American Ones

The results of acoustical anlysis of American vowels by Potter and Steinberg are plotted on the same f_1-f_2 plane of Fig. 6 and presented in Fig. 10 for the sake of comparison. The following points are to be remembered; that the Potter-Steinberg's data are based upon the steady part from the naturally spoken c-v-c syllable sounds and that there is no pitch test within a single speaker.

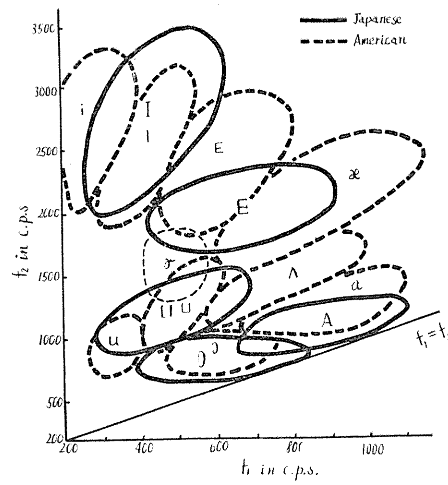


FIG. 10. Comparison of phoneme position of Japanese vowels with American ones.

Discussion

If we persuade ourselves of existence both of vocal quality and phonal (or phoneme) quality in our vocalic voices, we had better discard the element of vocal quality in order to display the phoneme figure for the best. This means the purest representation of phoneme quality in spite of the difference of vocal quality. As we have pointed out already, the separation of these two qualities is difficult, especially in the lower formant region in voiced vowel. We cannot say consequently that the differentiation of phoneme quality from vocal quality was quite perfect in this treatment. Here we have reported that we were aware of the existence of one- and two-formats in vowels "A" and "O." We are not sure whether this phenomenon can be attributed to the influence of the so-called invariant or of some other factors. In such a confused case we managed to average the formant position around the prominent component by weighting method as mentioned above.

This method of two-dimensional representation consists in outweighing the two formants picked out as the representatives of phoneme expression. We do not know whether what is meant by the third formant must deserve the treatment of essential formant for phoneme expression or must be discarded as invariant.

At any rate, this representation is not perfect in the sense that it can show only the heights of formant and not their intensities. And we do not think the perfect representation of phoneme figure can be expressed merely by the index of formant height. But even by adopting such a simple method we can show the characteristic phase of phoneme configuration of some vowels that are familiar in our language.

It is needless to say that a comparison of Japanese vowels with American ones is carried out only for the sake of simple convenience. For the results of Potter and Steinberg are analyses due to the so-called sonagram, employing the callers constituted of the considerable number of specialists in phonetics, but not adopting the process of pitch- and level-matching on the side of uttering subject as controlling technique. The vowels employed there are found quite natural in syllabic pronunciation. On the contrary, our results came from a stringent timbre study of vocalic voice which forces the uttering subjects to be constant as to their pitch and level of vowel pronunciation: as we have mentioned above the calling subjects of our experiment are all amateurs in phonetics. The vowels therein were conditioned in sustained state, because we aimed at a timbre study of vowel in static state.

Summary

When we take account of the effect of breathing intensity or uttering level, the voiced vowels cannot be easily treated; because the determination, and the separation of their variant and invariant formants are very difficult. In spite of the example of the lay caller employed here, the two-dimensional representation seems to be suitable to represent the phoneme figures of vowel. Through the Reports I and II, the following items are summarised;

(1) When the subjects, who are lays in phonetic or pronouncing technique, utter under the same condition of phoneme and pitch, there are still some amount of deviation of phoneme position.

(2) According to the pitch change, the phoneme position of a single vowel shows a considerable movement within the subject.

(3) Difference between individuals of phoneme position of vowel is large.

(4) Significance of (2) and (3) is left to be tested statistically, but it is noticed that both have some trends of deviation or movement.

(5) As cardinal American vowels, the vowel-square can be drawn, and for Japanese, the vowel-pentagon will be drawn.

(6) The change of this pentagon with pitch, level and individuality is left for

further investigation.

Acknowledgement

The authors are indebted to Mr. Masayoshi Nomura, Assistant Professor of the Department of Linguistics, Faculty of Literature, Nagoya University, for his valuable criticism and suggestion for this work. Thanks are expressed to Dr. Yoshiyuki Ochiai, Professor of the Faculty of Engineering, Nagoya University, for his copious advise giving a lead and also for his helpful discussion.