STUDY ON A PARTICULAR PROBLEM OF TIMBRE PATTERN OF ORAL VOWELS—CREVASSE PHENOMENA IN FORMANT STRUCTURE— DUE TO VOCAL FORCING IN PITCH AND LEVEL

YOSHIYUKI OCHIAI, TERUO FUKUMURA and TAKASHI SAKURAI

Department of Electrical Engineering

(Received October 31, 1962)

Summary—This is an advanced and detailed study of timbre pattern of oral vowels, treating especially pattern shape from peak and glen structures, and, at the same time, endeavoring to clarify the meaning of the so-called formant in its broadest meaning. Our study of timbre pattern begins with the selection of two calling subjects having contrasting nasal cavity conditions: one in perfect free state and the other in an inherent partially blocked state. Thus we can fully investigate and discuss the problem of perfectiveness and imperfectness of absorbing performance of nasal cavity in formation of oral vowels issued exclusively through buccal-cavity passage. Conspicuous contrast is found between the timbre patterns of our two callers, particularly when timbre patterns are obtained under the condition of level forcing which is accompanied by pitch change.

Introduction

In numerous examples of timbre patterns carried out in our timbre study, we found that for some types of voice the main formant of oral vowel "A"(ア) in particular is clearly of double-peak structure, and for other types of voice the main formant of the same vowel is of single-peak structure. This variation of structure caused by voice kind is significant. The similar variation can be seen in vowel "O"(オ). The double-peak structure usually occurring in ordinary Japanese vowels "A" and "O" reminds us of the so-called double-formant with short formant-interval. Notwithstanding this resemblance, they seem in no way similar. We feel it is necessary to bring out the difference between the double-formant with short interval and the double-peak structure in single formant which appear to be similar despite an essentially different nature, because in that way we are given a clue for a perfect explanation of formant construction.

There are other important phenomena in which, under certain conditions, the double-peak structure changes into a single-peak and, conversely, the single-peak structure changes into a double-peak. This is clearly revealed only in the timbre study of forced vocalics. Inasmuch as we should cover this very subject thoroughly, we must carry on a supplementary and advanced study of orals by utilizing the subtle change in pattern due to forcing pitch and level. This is the report of the studies carried out in our Laboratory in 1955.

Experimental Procedure

As our calling subjects, we use two young men: One, subject TF (abbreviated F) 30 years old, whose voice is quite familiar to us through numerous studies; the other, subject TS (abbreviated S) 23 years old, whose voice before testing seemed to offer considerable contrast. These two callers have approximately the same physique and stature. After testing, we inferred from the pattern study that the difference in their voice quality depended on the state and condition of the nasal cavity system. Fig. 1 shows the output-level characteristics *vs.* uttering

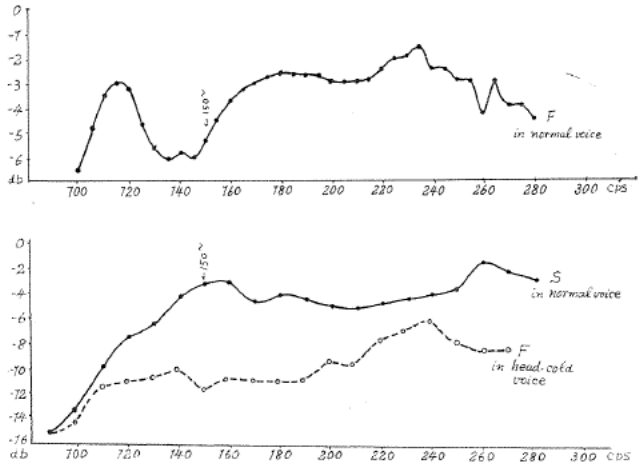

FIG. 1. Characteristics of output level of nasal sound [*Ng*] *vs.* pitch.

pitch of nasal /$\underline{\tilde{V}}$/* when the callers are instructed to maintain the uttering intensity at constant level of *mezzo-forte* with the help of their subjective judgement. From the representation of Fig. 1, we can quickly see that there is an inverse tendency between output characteristics of "F" and "S". For nasal output characteristic of the caller "F", the characteristic shows in the narrow region within 135 to 145 cps a very marked tendency—notwithstanding the subject's effort to keep a constant intensity of voice—toward a rapid and abrupt output decreasing, giving thereby an antiresonance-like curve form. Further, we can find another output decreasing at the point of 260 cps. In addition to the antiresonance-like tendency of these two points, we can find some resonance-like tendency at the points about 120 cps and 240 cps respectively (having a mutual relation of one octave) which results in a relatively large output production in spite of the skill of the subject and his effort to maintain level-constancy. These tendencies are nasal characteristics which are worth noting. For the voice "S", nasal characteristic shows no complicated form as does the voice "F", giving

* For the sound kept out of the mouth by being sent through the nose when the tongue is forced against the palate, we use the symbol /$\underline{\tilde{V}}$/. (refer Bibl. 7, 8, 9 .)

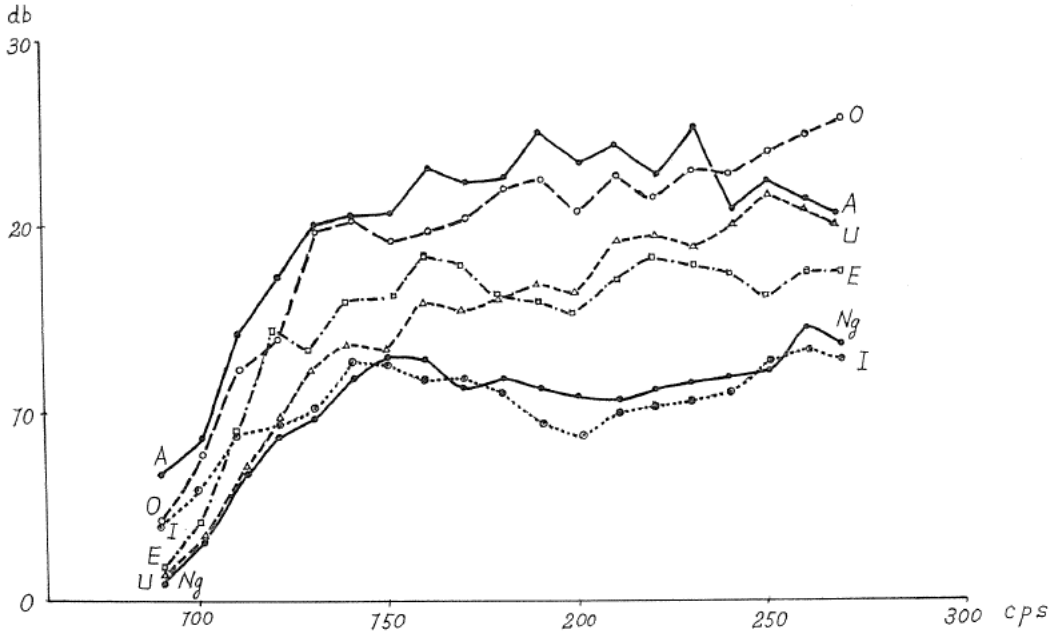

FIG. 2 (a). Voicing area of caller S.

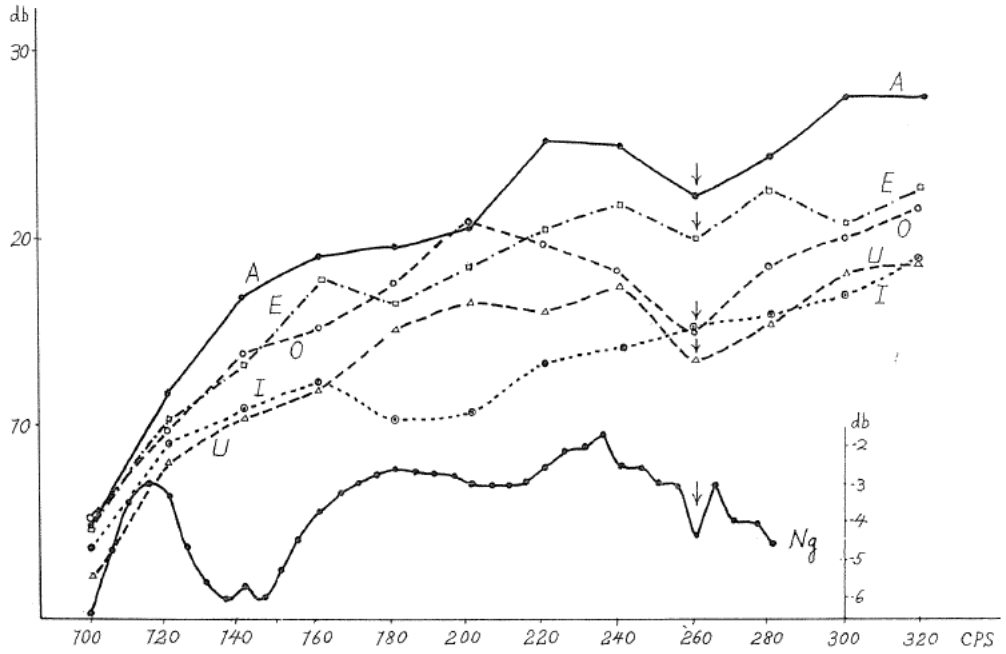

FIG. 2 (b). Voicing area of caller F.

only a monotonous flattened form. For reference, we show in the same figure the identical nasal output characteristic of the caller "F" when his nasal passage is in a condition of head cold, and we can see that both resonance-like and antiresonance-like characters are almost lost, resulting in a somewhat flattened characteristic which resembles the characteristic of the caller "S" in his normal state. In Fig. 2, we show output-level characteristics of five Japanese oral vowels for the callers "F" and "S" under the uttering condition of constant *mezzo-forte*. Fig. 2 (a) indicates the characteristics of the caller "S" and Fig. 2 (b) the characteristics of the caller "F". In both figures, we have added one more characteristic of nasal /$\tilde{V}$/, *i.e.*, /*Ng*/ described above. By comparing oral and nasal characteristics, we can easily point out a most interesting fact: for the voice "S", charatceristic of "I" vowel is quite similar to that of nasal, and for the voice "F" the output-reluctant position (260 cps) in vowels "A", "O", "E", "U" can be explained by the point where the nasal characteristic is abruptly repressed. It must be noted that for the voice "F" there is no oral vowel such as found in the voice "S" having a characteristic akin to nasal characteristic.

For pattern determination, eight pitches 130, 140, 150, 160, 170, 180, 190, 200 cps, with a constant difference of 10 cps, are used. As to level forcing, we use three stages, that is, *piano*, *mezzo-forte* and *forte*, with a level difference of constant 6 db between *piano* and *mezzo-forte* and between *mezzo-forte* and *forte*. It goes without saying that there is no strict way to maintain the

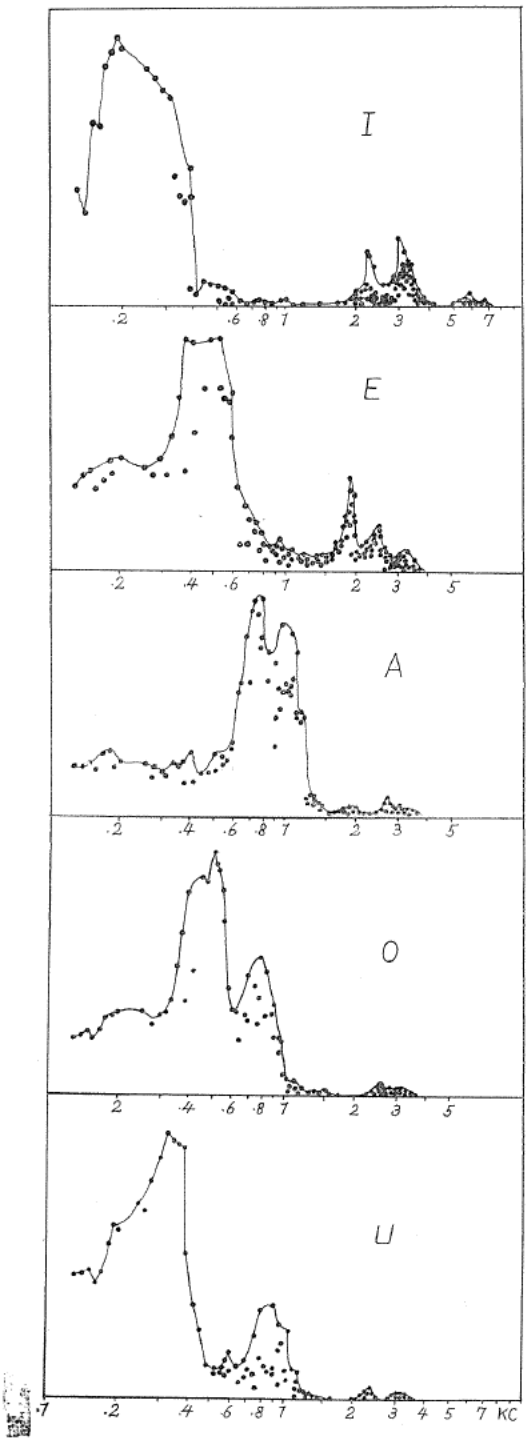

FIG. 3. Full patterns of five vowels of subject S.

constancy of uttering level, because there are no other means to supervise constancy except by employment of VU-Meter placed in front of the calling subject.

Experimental Result

To obtain the best timbre pattern with a view to clarify every point up to the fine structure, prudence in selection of uttering pitch is essential. For selection of pitches, we utilize the output characteristics as shown in Fig. 2. Eight pitches of small constant difference are best fitted for this purpose. We adopted the pattern-obtaining method and found it highly satisfactory, because, by it, we can trace the finest phases of pattern construction with ease and confidence. As a first step, we study phonemic patterns of five Japanese vowels obtained with the care the method requires. In Fig. 3 we list five vowels; in Fig. 3 we give phonemic patterns of the caller "S". These vowels are uttered at eight pitches on the intensity level of *mezzo-forte*. By regarding attentively Fig. 3, we can point out that in vowel "A" there is a shallow valley (dell) in the top part of the formant (at about 0.8 kc), and that in vowel "O" there is a relatively deep valley (glen) situated at about 0.6 kc. Next, we must study the change in

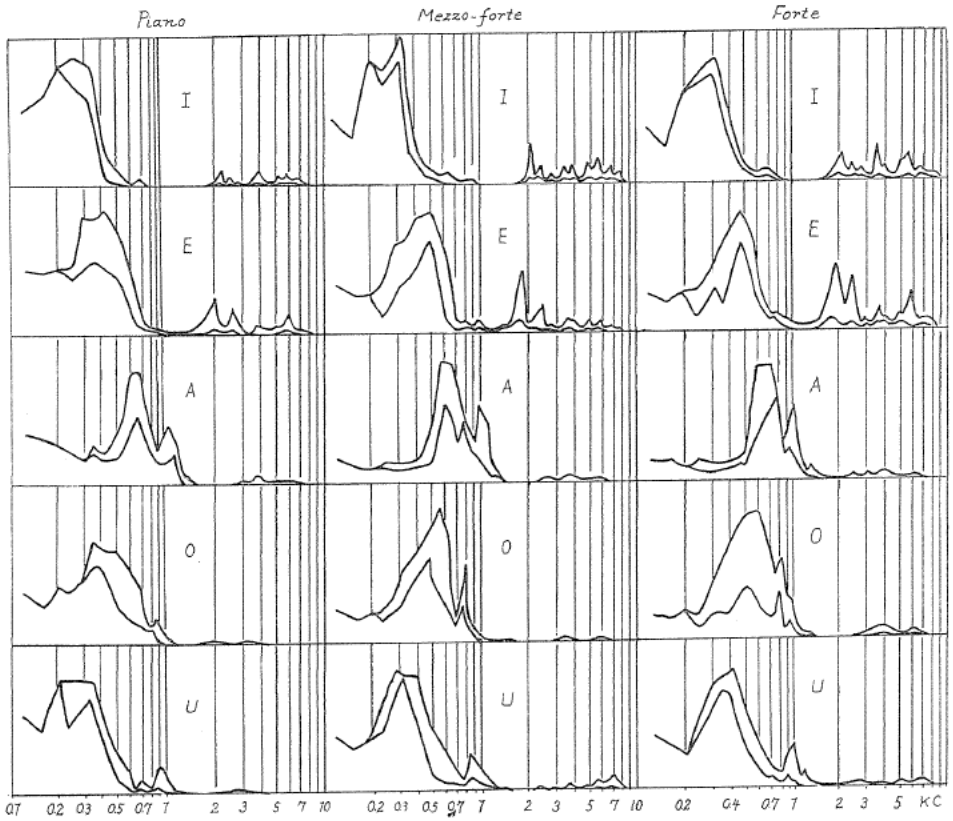

FIG. 4. Timbre patterns of five vowels uttered by subject F for three stages of vocal intensity: *piano, mezzo-forte, forte.*

pattern caused by change in uttering level. In Fig. 4 we give phonemic-pattern representations of five vowels pronounced by the subject "F" at three stages of different uttering intensities, that is, at *piano, mezzo-forte, forte.* The level difference between *piano* and *mezzo-forte* is about 6 db and is nearly equal to the difference between *mezzo-forte* and *forte*; the difference between *piano* and *forte* is about 12 db. The patterns given in Fig. 4 are advantageously displayed by two envelops, the upper and lower, by which we can easily divine the degree of standard deviation throughout the whole patterns. Another advantage of representation by two envelops is that we can make a further and closer examination of the pattern construction, for example, a particularized study of peak structure and glen structure. We can trace the exact position of peak and glen not only from the upper envelop but from the lower envelop as well and can study the details of pattern construction of peak and glen character.

In order to perfect a pattern study, we dash forward to the problem of *crevasse* in the formant-mountain found in Japanese vowels "A" and "O" which must be differentiated from the so-called double-formant with short interval. Essentially different from double-formant structure, which is indisputably composed of two distinct separate formants, in the notched formant we sometimes encounter cases where a compact mass of any one formant structure has some exceedingly narrow portion abruptly cut out. For instance, in Fig. 4 we can point out some examples, such as for "I" in *mezzo-forte*, for "A" in all three stages of level, and for "O" in *mezzo-forte*, all belonging to the same category of phenomena which we call formant-crevasse. Consequently, we think the term "crevasse in formant" is most appropriately applicable to the formant of Japanese vowels "A" and "O". If the formant of "A" and "O" could be considered double-formant with short interval, then it would be natural to consider that "I" vowel in *mezzo-forte* must be triple-formant. It is illogical to think that vowel "I" is double-formant in both *piano* and *forte*, and that it becomes triple-formant in *mezzo-forte* only. It is also unreasonable to consider that the so-called formant structure, *viz.*, the phonemic value, depends on the level of uttering force. We think that notched formant, as it appears, depending on the state of level and pitch forcing, has nothing to do with the phonemic value in phonetic aspect.

Now we return to our vowels "A" and "O". For the voice "F", formant-notching always occurs in vowel "A" throughout three stages of uttering level; in vowel "O", formant-notching occurs distinctly in *mezzo-forte* and less so in *forte* and is no longer clear in *piano.* That formant-notching is unstable and depends to some extent on the condition of level forcing, is very suggestive. But before entering in the conclusion of its causes. let us continue our observation a little more. In Fig. 5 we show phonemic patterns of vowel "A" for individual pitches. In Fig. 5 (a) we give phonemic patterns of the caller "F" and in Fig. 5 (b) those of the caller "S". For "F" voice, we show the lower pattern by mean envelop representation for five pitches (130, 140, 150, 160, 170 cps), and for "S" voice we give also mean envelop representation for six pitches (140, 150, 160, 170, 180, 190 cps). All of these are phonemic patterns of vowels uttered at *mezzo-forte.* In the voice "F", "crevasse in formant" is observed for all pitches except in patterns for pitch 150 cps. This exception corresponds to one of the pitches which do give the same effect for the voice "S", as is clearly shown in Fig. 5 (b),

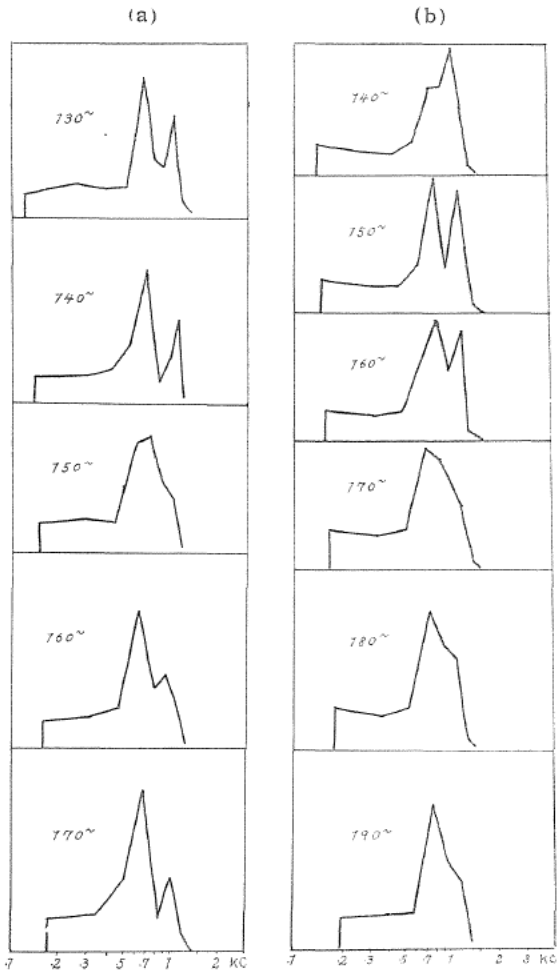

FIG. 5. Patterns (lower part only) of vowel "A", at vocal intensity of *mezzo-forte*, for several pitches: (a) for subject F; (b) for subject S.

where pitches other than 150 and 160 cps give formant with no crevasse. As a point of contrast, most of the pitches of the voice "F" give a formant with crevasse; on the contrary, the voice "S" gives formant without crevasse for most pitches and for a few only. it gives formant with crevasse.

It is worthy of mentioning that from the standpoint of nasal pronounciation (refer Fig. 1), the most important pitch of 150 cps (meaning for the caller F a unique case of no-crevasse, on the contrary, for the caller S one of the rare cases of violent crevasse) corresponds to one of the depressing pitches for F, but for S, to one of the most prevailing pitches.

Now we examine the effect of level forcing upon the formation of notched formant. In Fig. 6 we show phonemic patterns of vowel "A" uttered by the caller "S" on the same pitch of 150 cps and at three different stages of level, *viz., piano, mezzo-forte, forte.* In this representation, we can easily see that for

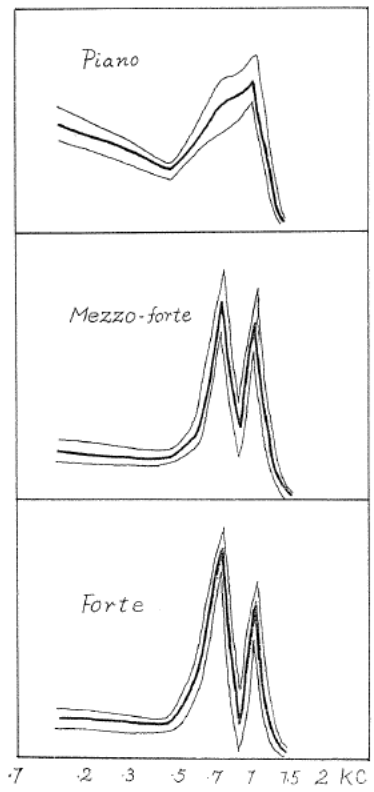

FIG. 6. Lower patterns of vowel "A" of subject S, for three stages of vocal intensity.

both *mezzo-forte* and *forte* the pattern shows a deep crevasse in the region 0.8–0.9 kc. and that for *piano*, the pattern shows no crevasse in this region; instead it reveals a little repressed area in the lower region of 0.4–0.5 kc. By these observations, we know well that notched formant, or formant with crevasse, in itself, is keenly sensitive not only to pitch forcing but also to level forcing.

Finally, we should like to add one more point most important in the sense that it clarifies the particular trait which distinguishes notched formant from ordinary double-formant, as clearly seen by the representation of Fig. 6, that formant-notching usually occurs as a sudden repression of one component only; at most,occurring as repression of two components. Repression of the one component, or two, at most, forms an abrupt attenuation in the pattern.

Another side of pattern studies, that concerned with peak positions of formant structure as influenced by the change in uttering level, has already been studied by us and the results have been published in a report titled "Timbre Study of Forced Vocalics". Our conclusion is that the peak position of lower part of pattern is undoubtedly changed by level forcing and, further, that the position of peaks in upper part of pattern does not change, irrespective of level forcing. In Fig. 7 we give an overall representation of both peak and glen structures of timbre pattern by using patterns of five vowels of the caller "F". Here we show only two levels, *piano* and *forte*, omitting *mezzo-forte* because by using only extreme levels of *piano* and *forte* we can easily and clearly trace the position shifting in peak structure. Here we use the symbol ↑

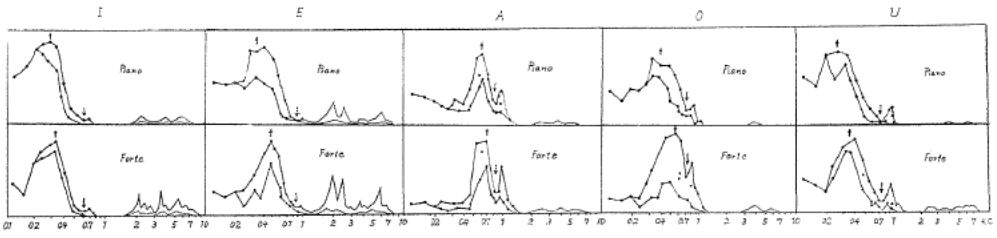

FIG. 7. Variation of timbre patterns of five vowels for subject F, caused by difference of vocal intensity: *piano* and *forte*.

to indicate the fundamental peak point and the symbol ↓ to indicate the fundamental glen position.

Discussion

We have investigated here the problem of formant division in vowels "A"(ア) and "O"(オ) particularly under the several conditions of pitch and level, by using exclusively two callers who stand in big contrast with regard to nasal performance. After careful inspection of this phenomenon, we are inclined to suppose that there must be an absorbing action of some acoustical system participating in the pronounciation of these vowels and thus bringing about a dividing effect in their principal formant. As one of such systems, a nasal cavity comes on the surface. Considering many aspects of this phenomenon brought into the light by this experiment, this way of thinking seems, to some extent, valid. If the crevasse phenomena in Japanese vowels "A" and "O" can be attributed to some other effect than this, the crevasse may be found in other vowels than these two. But in the system of Japanese vowels, there are only five vowels, and as open vowels having powerful formant in the vicinity of 1000 cps region we have only two, "A" and "O". In order to find similar phenomena, we must search other vowels in other languages.

As it may be, the problem remains as to the number of formant. Limiting to the Japanese vowels, which is better to decide: one formant or two formants concerning the vowels "A", "O"? From the physical viewpoint, envelop of formant becomes unique and single, for certain pitches, and for other pitches, is divided in two: according to the pitches, the same phoneme has single formant in some cases and double formant in other case. Phenomenologically, it is true. But, from the standpoint of phonemic value, the single-formant "A" is completely equivalent to the double-peak formant "A". In the discussion of formant, it is always important to consider not only the physical aspect but also psychological one.

Bibliographies

1) Y. Ochiai and T. Fukumura, Timbre Study of Vocalic Voices. MFE (Memoirs of the Faculty of Engineering), Nagoya Univ., Vol. **5**, No. 2 (1953).
2) Y. Ochiai and T. Fukumura, Timbre Study of Vocalic Voices Viewed from Subjective Phonal Aspect. Part. I. MFE, Nagoya Univ., Vol. 8, No. 1 (1956).
3) Y. Ochiai and T. Fukumura, Timbre Study of Vocalic Voices Viewed from Subjective Phonal Aspect. Part. II. (a). MFE, Nagoya Univ., Vol. 8, No. 1 (1956).
4) Y. Ochiai and T. Fukumura, Timbre Study of Vocalic Voices Viewed from Subjective Phonal Aspect. Part. II (b). MFE, Nagoya Univ., Vol. 8, No. 2(1956).
5) Y. Ochiai and T. Fukumura, Timbre Study of Vocalic Voices Viewed from Subjective Phonal Aspect. Part. III. MFE, Nagoya Univ., Vol. 8, No. 2 (1956).
6) Y. Ochiai, Memoirs on Nasalics. MFE, Nagoya Univ., Vol. **9**, No. 1 (1957).
7) Y. Ochiai and T. Fukumura, Timbre Study on Nasalics. Part. I. MFE, Nagoya Univ., Vol. **9**, No. 1 (1957).
8) Y. Ochiai, T. Fukumura and K. Nakatani, Timbre Study on Nasalics. Part. II. MFE, Nagoya Univ., Vol. **9**, No. 1 (1957).
9) Y. Ochiai and T. Fukumura, Timbre Study on Nasalics. Part. III. MFE, Nagoya Univ.

Vol, **9**, No. 2 (1957).
10) Y. Ochiai and T. Fukumura, Subjective Quality Measurement of Quality Characteristics Japanese Vowel "U", (Unpublished).
11) Y. Ochiai, T. Fukumura and T. Ina, Artificial Nasalics Obtained by Processing Five Oral Japanese Vowels, (Unpublished).
12) Y. Ochiai, T. Fukumura and T. Ina, Pattern Study of Artificial and Natural Nasals Inclusive of Interpretative Study of Timbre Patterns of Ordinary Oral-Vocalics Viewed from Recurrent Absorption due to Nasal Tubular Cavity, (Unpublished).