TIMBRE PATTERN REPRESENTATION AND SUBJECTIVE QUALITY MEASUREMENT OF "FORCED VOCALICS"

YOSHIYUKI OCHIAI and TERUO FUKUMURA

Department of Electrotechnics

(Received 31 May, 1960)

Synopsis—Following the basic line of thinking already described in our previous report "Introductory to Timbre Study of Forced Vocalics", we here present the experimental data of physical timbre-pattern study and of subjective quality measurement of forced vocalics. In accordance with our own concept of quality, *i.e.*, dual quality aspect, we give two kinds of pattern representation, that is, phonemic pattern and vocal pattern, and we present two kinds of fundamental quality, that is, articulation and naturalness, which are most important for speech communication. In addition to these fundamental qualities, we introduce here one other quality more primitive and direct in timbre judgement and examine the nature of this third quality of *piano-forte* timbre discrimination. Based upon the experimental results of pattern variation and quality difference of oral vocalics due to level forcing, we try to interpret the mechanism of quality formation in timbre pattern and to re-examine the concept of so-called *formant*.

Introduction

In our "Introductory to Timbre Study of Forced Vocalics",[1] we first viewed the several standpoints from which the forcing problem could be treated. For example, the subject of FORCED VOCALICS, including both level forcing and pitch forcing, is the one which touches upon accent and intonation in phonetic meaning. Considered from the standpoint of fundamental qualities of vocalics in speech communication,[2][3][4] this subject is not only important to phoneticians but is of great concern to communication engineers, because by studying this forcing problem we are necessarily led to the classification of communicational information, that is, the differentiation between phonemic-type information and vocal-type information. For the convenience of our description, we give first the experimental data of our pattern study[5][6] which was carried out in 1952 in our Laboratory and then present the result of our subjective study executed in 1955 and not yet published.

I. PATTERN STUDY OF FORCED VOCALICS

In our previous paper "Timbre Study of Vocalic Voices",[7] we introduced the method giving the phonemic pattern as well as the vocal pattern. Timbre patterns of vocalics including phonemic and vocal are sensitively affected by a change in forcing. To illustrate the forcing effect on timbre patterns, we give here only one example of our best caller whose voice and vowel has become familiar to us through many studies.

In Fig. 1, we show the so-called "voicing area" diagram and the "dynamic range" diagram for five vowels for this calling subject. Voicing area diagram means output-level of vocal egress *vs.* uttering pitch characteristics* and the dynamic range means the level difference of pronounced vowels between stages of uttering intensity, As usual in oral vocalics, there is no significant gain in discussing voicing area and dynamic range of vocalics, but if, in this case, there is anything to say, it is on the slight output-willing characteristic tendency of uttering pitch 240 cps and perhaps also on the output-yielding feature of vowel "A". Among five vowels, "A" is thus observed to be the one which has the maximum output level; "E" and "O" come next, having nearly equal output levels, and finally come vowels "I" and "U" as of equally poor output levels. It is also observed that the line of *mezzo-forte* runs not exactly in the middle between the lines of *forte* and *piano.* The mean average of the difference in VU-level between *forte* and *mezzo-forte* is about 6.0 db and the mean difference in level between *mezzo-forte* and *piano* is about 6.0 db. To detect the delicate and complicated nature of the change in pattern due to forcing, we must make pitch selection for pattern determination with deliberation. We use six pitches 120, 160, 200, 240, 280, 320 cps, having constant difference of 40 cps. In such a level arrangement and pitch selection purposefully and deliberately planned, we obtain the timbre patterns of both phonemic pattern and vocal pattern. In Fig. 2, we show the vocal patterns when they are influenced by the change in level and pitch. In Fig. 3, we show the phonemic patterns when they undergo the influence of level and pitch forcings at the same time. As we have described fully in our previous paper, vocal pattern representation is most important for expressing and revealing the so-called voice shading of individual persons, and meanwhile the phonemic pattern representation is most significant for representing vowel nuance of individual phonemes. It goes without saying that "voice" and "phoneme" are both concepts abstracted from idealized consideration and that "vocal pattern" and "phonemic pattern" are embodiments of these concepts in timbre aspect.

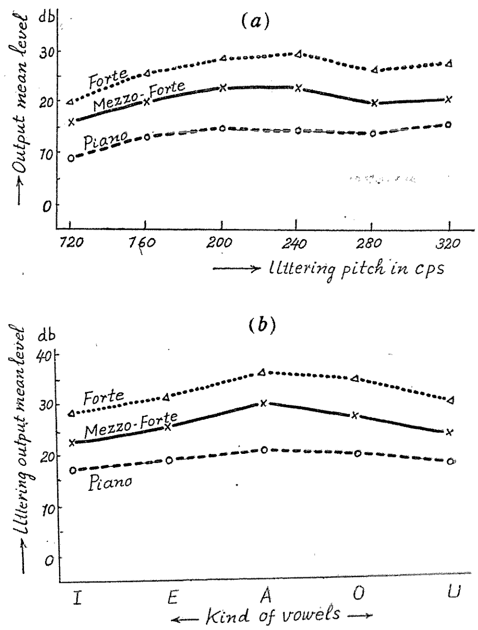

FIG. 1. Representation of "voicing area" characteristic for three stages (*f*, *m*, *p*) of uttering intensity of calling subject "TF" and "dynamic range" diagram for his five oral vowels (I, E, A, O, U).

* We will explain this in more detail later.

Change in Vocal Pattern Due to Forcing

From the vocal pattern representation in Fig. 2, we can trace the pattern change caused by the change in levels of two stages (*forte, piano*) for each of six different pitches (120, 160, 200, 240, 280, 320 cps). An inspection of these pattern changes teaches us that there is clearly a change in the lower pattern as a

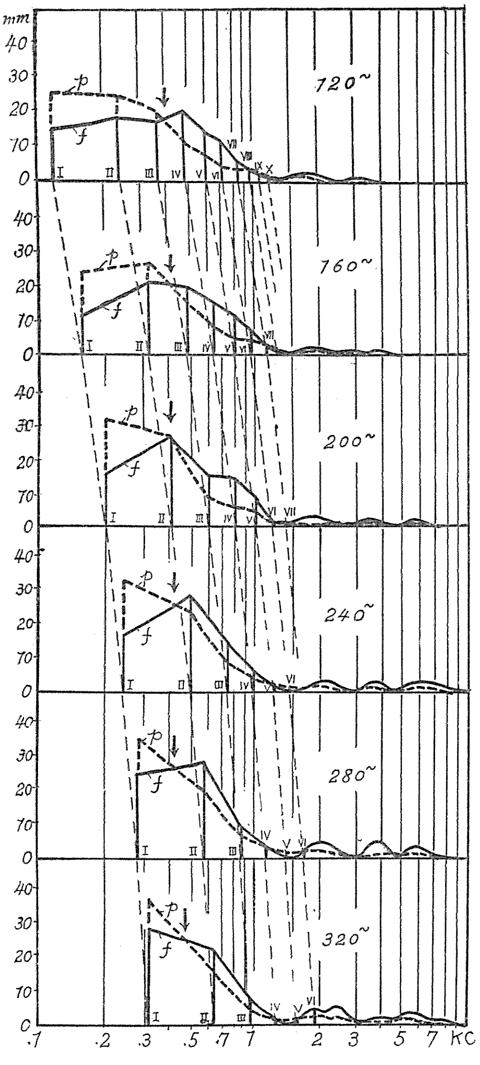

FIG. 2

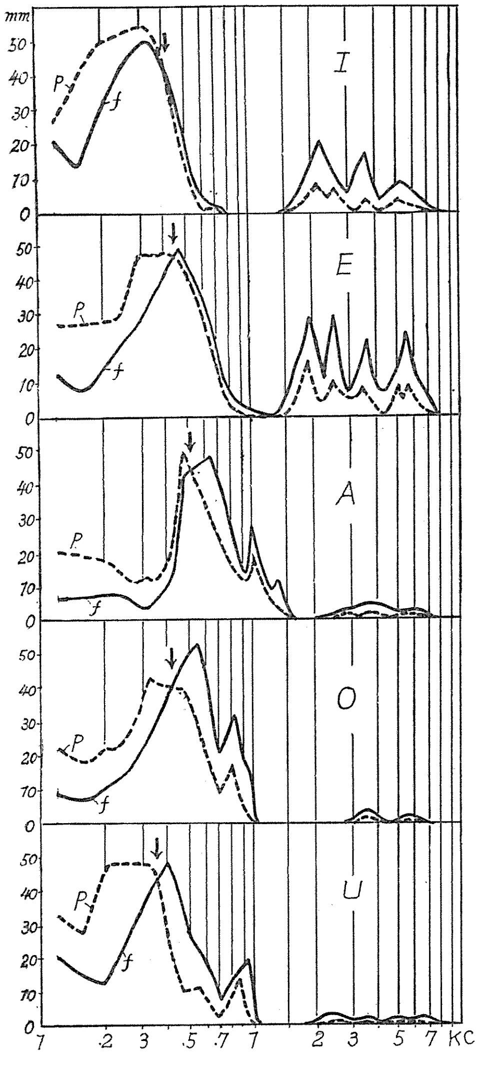

FIG. 3

FIG. 2. Change in vocal patterns of calling subject "TF" for his six uttering pitches, caused by the difference in uttering force (*forte* and *piano*). Marks ↓ show the main intersection point of vocal patterns by level-forcing of about 12 db.

FIG. 3. Change in phonemic patterns of calling subject "TF" for his five oral vowels, caused by the variation in uttering force (*piano* and *forte*). Marks ↓ show the main intersection point of phonemic patterns by level-forcing of about 12 db.

whole, *i.e.*, in the part of the pattern below the so-called vocal glen of 1.3 kc and that there is also a change in the upper part of the pattern lying beyond 1.3 kc which is unclear in this representation of vocal pattern. Accordingly as the uttering level and voicing pitch are increased respectively and independently, the pattern is changed in such a way that the amplitude of fundamental pitch is increased, accompanied by a simultaneous change in which the higher-order components are decreased. But, on the other hand, we cannot overlook the importance of the tendency toward establishment of the predominancy of even-number harmonic components, such as II, IV, VI, etc. harmonics (especially in patterns for pitches 120, 160 and 200 cps), which tendency is spurred on, particularly when the uttering level is increased.

As to the change in the upper pattern, we find a more intensified envelope for *forte* than for *piano*, but by this vocal pattern representation we cannot come to a definite conclusion on the fine points of the pattern difference.

We must point to one more fact which seems especially interesting and which, incidentally, is most important for our timbre theory: The intersecting point in the two patterns of two stages of level usually fall into a single position which is nearly constant in frequency dimension, notwithstanding the shifting of patterns because of pitch raising. This point of intersection of two patterns is situated at about 400 cps. It is true that it corresponds almost to the position of the so-called *invariant formant*[7] of this calling subject. Probably the constant position of *invariant formant* has some considerable relation to the intersecting phenomena of patterns due to level-forcing. But we cannot rashly conclude that the pattern-intersecting phenomena are brought about as the sole result of *invariant formant.* We must recall and consider again the influence of the process of level matching by VU-Meter which we took as most convenient for our timbre representation, because we need to remember that the movement of the indicator of VU-Meter is determined exclusively by the distribution of the partials in the lower pattern and is decided particularly by the components in this region.* In any event, we must appreciate and take an especially serious view of the fact that the patterns intersect at a certain single point. As to the reason for the phenomena of pattern intersection, we consider, for the time being, not only the effect of *invariant formant* of the caller but also the effect of level-matching process, both of which, we think, can play the important roles in and are responsible for the phenomena in question.

Change in Phonemic Pattern Due to Forcing

When we turn our attention to the pattern change in the phonemic representation by referring to Fig. 3 where the detailed features of pattern change due to level-forcing is made clear not only in the lower pattern but to some measure in the upper pattern, we arrive at the following realization.

The change is most evident in lower pattern due to forcing and the mode of pattern change in this area is unerringly characterized by the deformation of

* According to the result of the most comprehensive and all-inclusive quality study of timbre patterns of vocalics recently carried out in our Laboratory, the balancing point of VU-characteristics of vocalics in LCD and HCD, coincides exactly with the very point now in question.[12]

pattern: Patterns are deformed in such a way that the lower part situated below about 400 cps, corresponding to the pattern-intersection point, shifts downward in amplitude dimension, and the higher part situated beyond about 400 cps shifts upward in amplitude dimension accordingly as levels increase.

We must attach importance to the fact that, as seen in the majority of cases of phonemic lower patterns, there is also a point of intersection of about 400 cps which is unique in having the characteristic of an invariable pattern position. There may be some cases where the two pattern-envelope curves corresponding to the two stages of level come in contact such as with vowels "I" and "A" where the abruptly falling or rising parts of steep cliffs (front cliff in "A" and back cliff in "I") correspond to this critical region and consequently we cannot determine the exact position of pattern intersection. As to the phonemic higher pattern beyond 1.3 kc, there is a distinct change in mode, that is, by level forcing, the pattern is changed only in stronger direction, and by lowering the level, the pattern is changed only in weaker direction. In other words, in the area of upper pattern, the change in pattern seems to come only in the vertical direction of amplitude dimension and not in either upward or downward direction (of frequency dimension), that is, it never comes in the horizontal direction of frequency dimension. This is the distinct and significant difference in mode of pattern change between lower pattern and higher pattern.

Change in White-Phone Pattern Due to Forcing

We think the most remarkable trend of pattern change which holds good insofar as the voices uttered and vowels pronounced are without fault, is best represented in white-phone display in Fig. 4. This figure shows an ultimate overall mode of pattern change of forced vocalics. As can be clearly seen in this white-phone representation, the mode of pattern change in lower structure is characterized by one point intersection of about 400 cps which seems to depend on the position of *invariant formant* of caller and on the process of level-matching as well, and the mode of pattern change in upper structure is marked by the movement in vertical direction only.

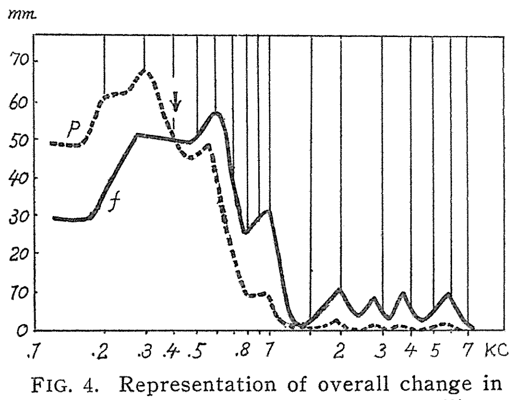

FIG. 4. Representation of overall change in all-inclusive vocal pattern display (W) for calling subject "TF", caused by the difference in uttering level (*piano* and *forte*). Mark ↓ shows the pattern intersecting point by level-forcing of about 12 db.

This is a résumé of pattern study of forced vocalics. As was clearly shown, the timbre patterns of vocalic sounds are sensitively affected by a change in uttering intensity. But to know what is, in fact, meant by these pattern changes due to forcing, we must go on to study the so-called subjective quality measurement because we are convinced that this subject of forced vocalics is most intimately connected with the very problem of how to recognize the principal and fundamental meaning of the so-called *formant*. In other words, we think that a quality study based on subjective measurement must have priority over the objective physical study. In a moderate way, we feel that the physical study alone cannot lead to any "determinism" of the meaning of *formant*.

II. QUALITY STUDY OF FORCED VOCALICS

Concepts of Fundamental Qualities and Nature of Subjective Measurement

Our direct aim in this subjective study is first to obtain subjective data of quality measurement and next to study the interrelation between subjective data on quality judgement and physical data on timbre pattern representation. It was planned that timbre representation of signal stimulus is to be made according to the fundamental quality aspect; two representations of pattern, one of vocal and one of phonemic, were prepared in the preceding section. In this subjective study, we must prepare to test at least two kinds of quality, *viz.*, the quality corresponding to phonemic pattern representation and the quality corresponding to vocal pattern representation. For *phonemic quality*, *viz.*, the quality corresponding to phonemic pattern, we must prepare some quality measurement touching on "phonemic values" of vocalic sounds, and the so-called *articulation* measurement serves our purpose. For *vocal quality*, *viz.*, the quality which corresponds to vocal pattern, we must seek for quality measurement of some other kind, one which most uniquely estimates the "vocal values" or "voice values" of vocalic sounds. We think that what we call *naturalness* measurement meets our need completely.

Identification and discrimination in pure phonemic aspect without any regard to the person to whom the vocalic presentation belongs, might give us the data of quality judgement by stressing only the phonemic side of vocalic perception. Identification and discrimination in pure vocal aspect, taking no notice of the phoneme of the vocalic to which presentation in fact corresponds, could give us the data of quality judgement by attaching importance only to the vocal side of vocalic perception. In other words: Articulation measurement made in pure phonemic aspect, uncontaminated by vocal attention, is most fittingly used for phonemic quality measurement. Naturalness measurement made in pure vocal aspect, not blurred by the phonemic distraction, is most appropriate for vocal quality measurement. Quality order in its conceptional consideration is primarily valuable for our quality theory before we go on to concrete and detailed measurement of quality, because in quality study we cannot over-exaggerate the need to make clear the objective and substance of quality itself with which we concern ourselves most profoundly.

Digressing, for the time being, from the quality theory viewpoint of scientific interset, let us recall our daily-life experience in connection with the forcing phenomena. It is true that voices which are too loud or too soft are both too inarticulate for our understanding, and that extremely high-pitched voices also lose, in some measure, their articulation and clearness. Our particular interest will be aroused by the fact that the extremely high-pitched voices as well as those too loud and too soft lose their vocal character to such extent that very often we err in judgement as to the speaker when he is out of sight. Phonemic values can sometimes become lost by forcing of level and pitch, but it is doubtless on vocal values rather than on phonemic values that the forcing phenomena produce the most vigorous effect. Considered from the essential nature of this forcing phenomena, our concern about vocal quality remains, and interest is thus directed toward vocal quality rather than toward phonemic quality. What further intensifies our purpose to stress vocal quality is the certain fact that studies so far made, particularly those made recently in the United States[8] and Soviet Russia,[9] are quality studies viewed from intelligibility aspect which does not cover the other side of quality, and this, in our opinion, is most vital and important to a thorough study.

Experimental Procedure

The steps in our subjective study of forced vocalics are as follows: To obtain the articulation characteristics in phonemic aspect judgement, we use a series of phonemes composed of five Japanese oral vowels pronounced in the ordinary way, and to get the naturalness characteristics in vocal aspect judgement, we use a series of voices taken from a group of five calling subjects composed of three young men and two young women.

To determine the most suitable pitch and level, we insist that at the first stage of experiment the callers do not employ an extreme unnatural voice caused by excessive forcing which would impair and excessively disturb their own vocal quality.* For that purpose, we are obliged to make use of a so-called voicing-area diagram for each caller and for each vowel, as shown in Fig. 5.

The voicing-area diagram[10][11] may now be explained more in detail: Even when the voice of the caller is uttered in an effort to maintain a constant uttering level, the output-level of vocalics from his mouth is not constant but changes accordingly as the pitch varies. The output-level *vs.* pitch characteristics can be obtained for each vowel and for each stage of uttering intensity. For example, for each level of three stages, that is, *piano*, *mezzo-forte* and *forte*, we can get five output *vs.* pitch characteristics for the five vowels of one person. By assigning five vowels and three stages of uttering intensity, we can obtain a great many characteristics from which we can determine three approximate domains of voicing area: *Piano-*, *mezzoforte-*, *forte*-area. By inspecting Fig. 5, we can understand that both the pitch compass and dynamic range are different, depending on sex and physical characteristics of the caller: The dynamic range from *piano* to *forte* is as a whole narrow for female voices ("C", "S") and is relatively wide for male voices ("F", "K", "Y"). The fact that in certain callers' voices there are found some overlappings or gaps between areas must be attributed to the circumstance that all these callers are but amateurs in vocal singing and have never received

* For the advanced stage of experiment, this type of experiment, that is, the excessive forcing, is of course necessary.

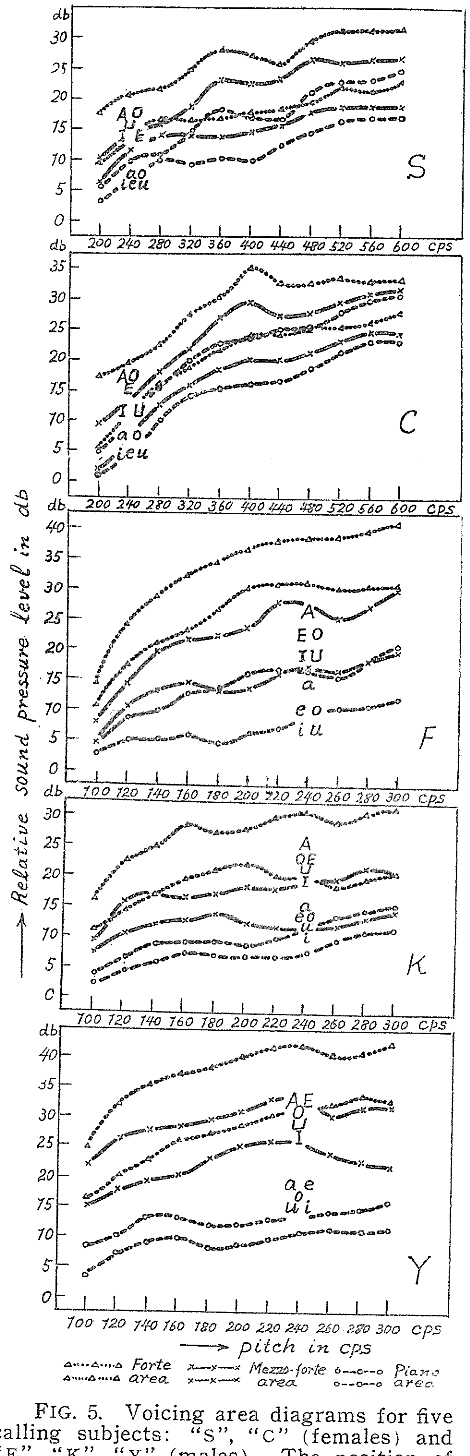

FIG. 5. Voicing area diagrams for five calling subjects: "S", "C" (females) and "F", "K", "Y" (males). The position of A, E, I, O, U in each figure shows the level and pitch of *forte* utterance and that of *a*, *e*, *i*, *o*, *u*, shows those of *piano* utterance.

any regular vocal lessons or training in singing. Our aim in this test on voicing area is, however, to find a suitable range and pitch for callers' utterance and not to decide the correct compass and right range, in the musical and vocal meaning. By considering the pitch compass and level range of individual callers, we determine the one pitch of 240 cps which is common to all, not too low for female callers and not too high for male callers. As to the level difference used for our subjective experiment, we can conveniently select two stages of level, *piano* and *forte*, the difference between which is about 10 db for female callers and about 15 db for male callers, which means that while for males the dynamic range is only partially used, for females the dynamic range is fully utilized. We made this choice because we think that a psychological measurement of this kind is usually poor in precision when compared with the precision of physical measurement such as found in pattern determination. This difference of 10–15 db in two levels assigned for our experiment is not so significantly large as to suit the real condition of *piano* and *forte* in vocal music, but we feel it is sufficient at the present stage of our experiment because we are now aiming at and interested in the study of forced vocalics having such a level difference inasmuch as we meet with it in ordinary conversation and not in vocal music.

In order to detect the influence of level-forcing upon timbre quality of oral vocalics from the psychological standpoint and, further, to interpret, if possible, the cue in the mechanism of building-up of fundamental qualities of vocalics in their pattern in frequency dimension, *i.e.*, to find the cue of quality formation in frequency dimension, we purposefully employ the distortion of band-eliminating type (BED) as most appropriate. With the aid of this distortion, we can determine the concrete change in quality *vs.* frequency characteristics due to level-forcing and, further, we might bring it into association with the change in timbre pattern in frequency dimension due to the same level-forcing.

In this experiment, we use MR-103 Condenser-Type Microphone and High-Quality Loudspeaker of 8-inch diaphragm. As an overall frequency characteristic of our experimental system, including tape-recorder, the range for the most reliable fidelity must be restricted to within 200–6,000 cps.

Experimental Result

As the most natural outcome of our carefully planned experiment, we obtain the quality change due to level-forcing by applying the distortion of band-elimination systematically and bilaterally. Of course, the naturalness and articulation qualities are considered and adopted as the most fundamental. In addition to these fundamental qualities, we try to make use of a third, which we temporarily term "timbre quality in *piano-forte* discrimination". By the nature of things, this third quality is of binary choice. The descriptive order of our experimental results is articulation, naturalness, and quality in *piano-forte* discrimination.

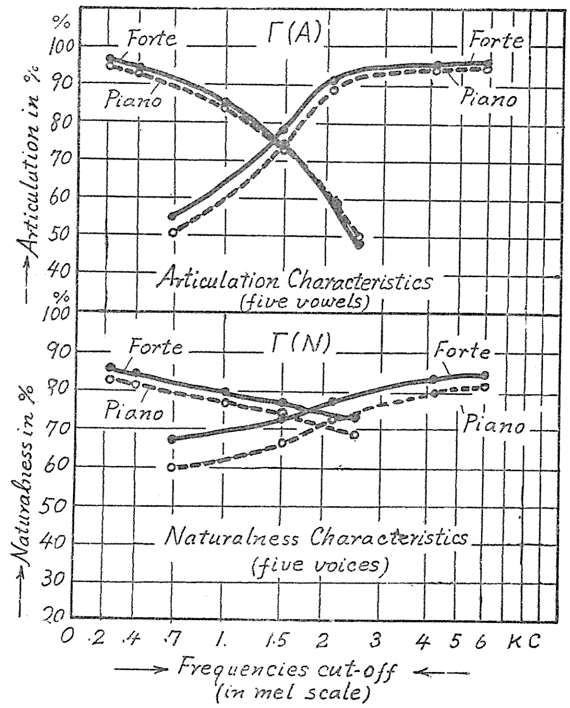

FIG. 6. Changes in articulation and naturalness characteristics caused by *piano* and *forte* expressions. Any one point of these characteristics is determined as the result of 1,500 observations by four listeners.

Articulation Characteristics $\Gamma(A)$

Articulation characteristics of level-forcing are shown in the upper section of Fig. 6. In this representation, full-line curves mean articulation characteristics in *forte* uttering-expression and broken-line curves indicate articulation characteristics in *piano* uttering-expression. Between the levels of *piano* and *forte*, there is 10–

15 db level difference as already mentioned. Articulation quality is measured in percentages and is expressed in linear scale on the ordinate. On the abscissa, the frequencies of band-eliminating cut-off are measured in "mel" scale. To determine any one point of these characteristics, we take a total of 1,500 observations by four listeners. As is clearly shown in this diagram, for characteristics in low-cut distortion (LCD) there is little or no difference between qualities in *piano* and *forte* expressions, meanwhile there is quite a noticeable difference between qualities in *piano* and *forte* expressions for the characteristics in high-cut distortion (HCD). It is consequently evident that there is little or no shift in position of balancing point of articulation quality caused by the level-forcing of about 12 db.

Naturalness Characteristics $\Gamma(N)$

Naturalness characteristics of forced vocalics are shown in the lower section of Fig. 6. For the determination of any one point of these characteristics, a total of 1,500 observations is prepared for each condition of distortion. Naturalness quality is measured in percentages in linear scale on the ordinate, and the frequency on the abscissa is measured in "mel" scale, exactly as under "*Articulation Characteristics*". The unimpared value at zero distortion with which naturalness characteristics start, is usually lower than that of articulation. This difference in starting quality must be attributed to the real condition that the degree of practice of listening crew is not the same for articulation and for naturalness. The gradient, *viz.*, the slope at the starting point of naturalness characteristics, is slighter, especially in LCD characteristics, than that of articulation characteristics, and, further, in naturalness quality, the characteristic slope does not change very much when the distortion is greatly increased. It is for this reason that the naturalness characteristics are seemingly strait and linear with slight inclination to horizon. This linearity and simplicity of naturalness characteristics are better represented by frequencies in "mel" scale than by frequencies in "logarithmic" scale. These are the well-known trends of naturalness characteristics in general, quite familiar to us.

As to the characteristic difference due to level-forcing which is the very objective of our study, we can find here a relation similar to that in articulation. That is to say, there is a relatively small difference in LCD characteristics and a relatively large difference in HCD characteristics.

As the result of this fact, we can see easily that there is a little shift in balancing point of naturalness quality caused by the change in level-forcing. This shift is about 50–100 cps directed upward by level-forcing of about 12 db.

By closely comparing the characteristics of articulation and of naturalness and by reflecting chiefly on, and extracting, the difference in quality response due to the change in uttering level, we can safely conclude that, whereas the articulation quality, including and summarizing both trends of characteristics in LCD and HCD, shows a relatively small difference on one hand, naturalness quality shows a relatively large difference on the other.

When we keep our eyes upon the balancing point of these characteristics, a slight difference can be found. Reviewing both articulation and naturalness, the balancing point of the characteristics in *piano* expression is shifted but slightly in position from the balancing point of the characteristics in *forte* expression. It is probably worth noting that there is quite a noticeable change in position of balancing point due to pitch-forcing* and that there is little or no shifting in position of balancing position due to level-forcing. If we now change the angle of our observation a little by first selecting a parameter of level of utterance, then we get another representation of quality characteristics, as shown in Fig. 7. Here we have two sets of fundamental quality characteristics for individual levels. A glance at such representations leads us to a quick understanding of the fact that in spite of the level difference, the fundamental quality characteristics are not affected and disturbed insofar as the form and movement of characteristics are concerned.

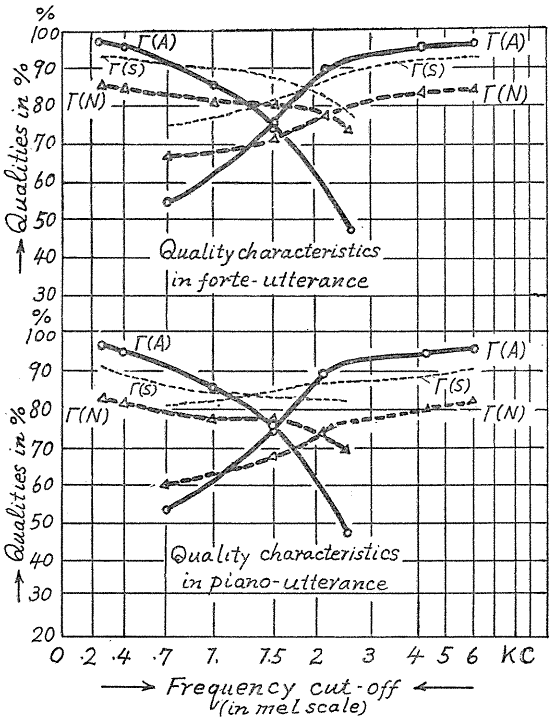

FIG. 7. Articulation, naturalness and timbre quality in stress discrimination in *forte* and *piano* utterances.

Quality Characteristics Γ(S) Based upon Piano-Forte Timbre Discrimination

It goes without saying that these quality characteristics are the results of the consideration and treatment of the timbre quality in pure sense. In other words, these quality results have little or no relevancy to the influence of loudness difference due to level-forcing.

As there is little or no participation of loudness clue in our quality judgement, the discrimination between *piano* and *forte* timbre signals must be related to timbre clue exclusively. It is for this reason that we introduced a novel quality based solely upon *piano-forte* timbre discrimination which has been most difficult yet most

* A report on this will appear later.

interesting. In addition to naturalness characteristics $\Gamma(N)$ and articulation characteristics $\Gamma(A)$ in Fig. 7, we give for reference and comparison the third quality characteristics, temporarily designating them by $\Gamma(S)$ and representing by light broken-line curves. For determination any one point on this third quality characteristics, a total of 1,500 observations is put in service for each condition of distortions. Through careful inspection of these three characteristics $\Gamma(A)$, $\Gamma(N)$, $\Gamma(S)$ in *piano* expression as well as those three in *forte* expression, we come to the realization that, notwithstanding the clearly evident difference in nature and content of judgement, there is, on one hand, some significant similarity in form between $\Gamma(N)$ and $\Gamma(S)$ and, on the other, there is considerable dissimilarity in form between $\Gamma(A)$ and $\Gamma(S)$. Then comes a question: Is, in fact, the timbre quality which is uniquely based upon judgement of *piano-forte* binary selection, exactly the same as the quality of naturalness which, in this case, is based exclusively upon judgement of choice of one among five voices? To answer this, we need further and more specialized observation. For this purpose, we give in Fig. 8 both *forte*-quality characteristics(solid-line curves) and *piano*-quality characteristics (broken-line curves) in the same graph. By carefully comparing the representation of naturalness characteristics given in Fig. 6, we find the following: In spite of the seeming similarity between $\Gamma(N)$ and $\Gamma(S)$, closer inspection reveals that there are fundamental differences between them which are slight in degree but of special importance. In naturalness characteristics $\Gamma(N)$, the quality corresponding to *forte* expression is always greater than the quality corresponding to *piano* expression.

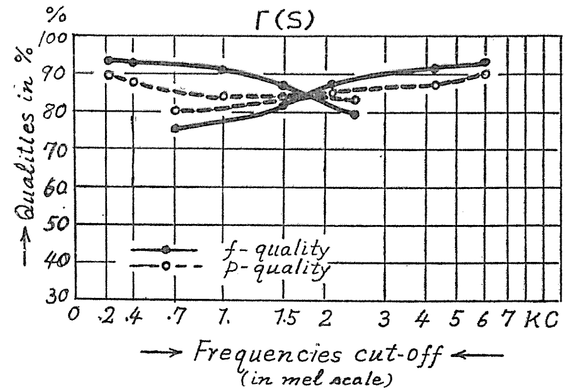

FIG. 8. Characteristics of *piano*- and *forte*-quality in *piano-forte* timbre discrimination. Any one point of these characteristics is determined as the result of 1,500 observations by four listeners.

The important feature which differentiates $\Gamma(N)$ from $\Gamma(S)$ consists not only in one-sided predominancy of naturalness quality in *forte* expression over that in *piano* expression but also in the fact that the quality difference between two expressions tends to increase—an increase not large, but slight—accordingly as the distortion increases. This increasing tendency of quality difference under increasing distortion is observed both in LCD and HCD and it is more evident in HCD than in LCD. This is the very point which distinguishes $\Gamma(N)$ from $\Gamma(S)$. In characteristics $\Gamma(S)$, on the contrary, we find the tendency toward crossing of two qualities, *forte* and *piano* qualities, when the distortion increases. In other words, in quality characteristics $\Gamma(S)$, the *forte*-quality is not always greater than *piano*-quality; the region where the *forte*-quality can be greater than the *piano*-quality is restricted only to the before-stage of distortion. After passing through the so-called balancing point, the quality characteristics, both LCD and HCD, undergo noticeable and significant changes and the *forte*-quality becomes inferior to the *piano*-quality. This means that the two qualities fall in reverse order after passing through the neighborhood of *vocal glen* of the patterns.

As the next point of importance to be found in characteristics $\Gamma(S)$, must point out the fact that the difference of characteristics between the two qualities, *piano* and *forte*, is great for LCD distortion and small for HCD. This unmistakable feature of $\Gamma(S)$ stands in fine contrast to other qualities, articulation quality $\Gamma(A)$ and naturalness quality $\Gamma(N)$.

III. INTERPRETATION OF EXPERIMENTAL RESULTS

After glancing at the résumé of pattern study of forced vocalics and reviewing the results of subjective measurement of quality-response characteristics due to forcing, we must now reflect upon the nature of their interrelation.

In the previous quality treatment of vocalics already published, our concern were exclusively with the finding of so-called correspondency between timbre pattern and timbre quality. For example, for the point demarcating pattern structure and also for the balancing point of quality distribution, we found it favorable to utilize the point of vocal glen. What we have so far clarified is: We infer that the vocal quality distributed in and allocated to the lower pattern is approximately equal to that distributed in the upper pattern. Further, from the information on slope of quality characteristics in BED, we can infer the degree of density of quality concentration in band section of pattern. Now that we take the parameter of level-forcing, the problem becomes more delicate and complex, because we must pose a question not only of the pattern form and quality value but also of the change in pattern form and change in quality value. We insist that our main interest in the present study consists of quality differentiation corresponding to pattern differentiation. This means that we here come in contact with the problem of higher-order quality than ever. Irrespective of whether we take phonemic pattern or vocal pattern, the effect of level-forcing upon pattern is characterized by the complicated tendency of the pattern toward intersecting in lower pattern, on one hand, and, on the other, by the simple modification of the upper pattern which is revealed as change in vertical direction only. When we view the nature of our problem by borrowing the terms of pattern information, what becomes of our problem? For example, in the following, we stress pattern-intersecting information. Stimulus of timbre signals of several different pattern-envelopes which intersect at a certain single point is introduced for our perception: What is the result? What is the difference in quality which responds to the difference in pattern in such an introduction? In other words: Is there any noticeable turning point in a set of quality characteristics to correspond with, for example, the crossing-over point in pattern?

As we have already seen, the pattern differences of *piano*- and *forte*-utterance correspond, in some measure, to the quality differences. There are, however, no turning points in naturalness characteristics $\Gamma(N)$ which correspond to the cross-over point in pattern; in contrast to this, turning points are found in timbre quality characteristics in *piano-forte* discrimination $\Gamma(S)$, but these turning points $\Gamma(S)$ are situated in the neighborhood of vocal glen and do not coincide with the cross-over point in pattern. As a matter of fact, we can find the reversible relation between two qualities, *forte*-quality and *piano*-quality, only in the neighborhood of vocal glen, and for this reason we cannot yet fully interpret this reversible relation.

As it may be, this point seems to require further discussion. We introduced the so-called phonemic pattern as revealing or characterizing or stressing the phonemic side of timbre pattern; but for the realization of phonemic quality, the entire pattern is not necessarily needed and only a few parts of the pattern which are the most essential of phonemic pattern are really indispensable. Reversely expressed, all envelope information of phonemic pattern is not necessary for formation of phonemic quality. For realization of naturalness quality, however, far wider parts of vocal pattern seem to be necessary. But, in this case also, naturalness itself will never come to need the information of the whole envelope of vocal pattern. We must stress only the fact that in comparison with the articulation, the naturalness needs far wider parts of pattern envelope. It seems to us for the formation of the third quality in *piano-forte* timbre discrimination that the widest part of pattern envelope is, in fact, needed. When we consider that the tendency of quality-turning phenomena of the so-called third quality must be related to the tendency of pattern crossing-over phenomena, that is, to the envelope-turning phenomena, then we can at least infer that the exceedingly wide range of pattern envelope must be needed for formation of this third quality. We can further infer that the more primitive and direct the timbre judgement becomes, the wider the envelope needed for its quality formation. Along the line of such thinking, naturalness is more primitive and direct than articulation, and the third quality in *piano-forte* timbre discrimination seems to be far more primitive and direct than naturalness quality.

Conclusion

The result of our research described in our subjective study of level-forcing Parts I and II can be concentrated in the following points of importance.

(1) As for quality difference due to level-forcing in band-eliminating distortion in the general meaning of fundamental qualities, including articulation and naturalness, the difference in LCD is always larger than that in HCD. When we translate this fact by using terms of quality formation in timbre frequency pattern, the change in quality distribution in the overall pattern due to level-forcing is relatively small in lower pattern and relatively large in upper pattern. As one of the more direct and primitive qualities, the so-called third quality is experimented with. This quality is obtained by simple timbre judgement of binary choice, for example, by judgement in *piano-forte* discrimination in this case. After reviewing and summarizing the characteristic trends of this quality, we can state with some deliberation, *viz*:

(2) The difference between characteristics of *piano* and *forte* qualities, *i.e.*, the

difference in two qualities corresponding to the pattern difference between *piano* and *forte* expressions, is more remarkable for low-cut distortion (LCD) than for high-cut distortion (HCD). In other words, the quality difference due to the change in level-forcing is more marked in the lower pattern than in the higher pattern, insofar as the distribution of quality difference is concerned. We can put it in another way: For the establishment and formation of the type of quality of binary selection, the pattern difference in the lower construction can have more significance than that in the higher construction.

(3) When we compare naturalness quality with articulation quality by referring to the effect of forcing, then we can find the order of quality change from the characteristic movements. The most conspicuous effect is found in naturalness characteristics subject to distortion HCD, and the least effect is found in articulation characteristics subject to distortion LCD. As to the two remaining characteristics, that is, to naturalness characteristic in LCD and articulation characteristic in HCD, the effect of level-forcing is moderate and less conspicuous. In other words, by raising the uttering level, quality formation in the pattern undergoes a change, *i.e.*, the most conspicuous is found in the upper pattern for naturalness quality and the least conspicuous in the lower pattern for articulation quality. Expressed reversely, by lowering the uttering level, the part of vocal quality attributable to the upper structure of vocal pattern is noticeably influenced and impaired, and the part of phonemic quality attributable to the lower structure of phonemic pattern is only very slightly influenced and deteriorated. The vocal quality allocated to lower structure of vocal pattern and the phonemic quality allocated to upper structure of phonemic pattern are both only moderately influenced by level-forcing.

The third quality related to binary selection of *piano-forte* discrimination is covered further in points (4) and (5) as follows:

(4) Timbre quality based upon *piano-forte* discrimination is seemingly similar to naturalness quality; but it in no way resembles articulation quality.

(5) The one point which distinguishes the so-called third quality from naturalness quality and which is equally as important as point (2), is that, whereas the *forte*-quality is superior to the *piano*-quality during a distortion range which is not too extensive, the *forte*-quality becomes so inferior to the *piano*-quality when the distortion range is increased that it surpasses the so-called vocal glen. This means: at the same time, taking into consideration both LCD and HCD characteristics, *forte*-quality is stronger than the *piano*-quality so long as the distortion remains moderate and weak, but *forte*-quality becomes weaker than *piano*-quality when the distortion becomes excessive and when it surpasses the so-called vocal glen.

Acknowledgement

This subjective study on forced vocalics has been made possible through a Grant-in-Aid for Fundamental Scientific Researches from Ministry of Education of Japan.

References

1) Y. Ochiai and T. Fukumura: Introductory to Timbre Study of Forced Vocalics. MFE (Memoirs of the Faculty of Engineering) of Nagoya University, Vol. **9**, No. 2, Nov., 1957.
2) Y. Ochiai: Transmission of Quality. MFE of Nagoya Univ., Vol, **6**, No. 2, Nov., 1954.
3) Y. Ochiai and T. Yamashita: On Timbre Quality. Part I—Introductory Consideration. MFE of Nagoya Univ., Vol. **7**, No. 1, May, 1955.
4) Y. Ochiai: Mémoire sur les Sons des Voix Humaines. MFE of Nagoya Univ., Vol. **4**, No. 1, July, 1952.
5) Y. Ochiai and T. Fukumura: Hassei-Kyodo ni yoru Onsei-Pattern no Henka ni tsuite (in Japanese)—On Timbre Change Caused by Uttering Force. Read at Joint Meeting of Societies of Elect. Eng., Commun. Eng., and Illum. Eng. of Japan, April, 1955.
6) Y. Ochiai and T. Fukumura: Hassei-Kyodo ni yoru Boin-Onshoku-Kozo no Henka ni tsuite (in Japanese)—On Change in Timbre Pattern of Vocalics Due to Uttering Force. Read at Meeting of Acoust. Society of Japan, May, 1955.
7) Y. Ochiai and T. Fukumura: Timbre Study of Vocalic Voices. MFE of Nagoya Univ., Vol. **5**, No. 2, Sept., 1953.
8) J. M. Pickett: Effect of Vocal Force on Intelligibility of Speech Sounds. JASA, Vol. **28**, No. 2, 1956.
9) М. А. Самсонов: Спектр Русской Речи и его Деформация под Воздействием Сильного Звукового Раздражитеря. Восприятие Звуковых Сигналов в Различных Акустических Условиях. ИЗДАТЕЛЬСТВО АКАДЕМИИ НАУК СССР 1956.
10) B. Stout: The Harmonic Structure of Vowel in Singing in Relation to Pitch and Intensity. JASA, Vol. **10**, No. 2, 1953.
11) T. Yamashita and M. Tsuda: Some Observations on Voicing Area in Vocalization Phenomena. MFE of Nagoya Univ., Vol. **6**, No. 1, May, 1954.
12) Y. Ochiai et M. Oda: Sur l'intensité sonore subjective des vocales soutenues ayant les significations phonémique et vocalique. MFE of Nagoya Univ., Vol. **11**, No. 1–2, November, 1959.