

INTRODUCTORY TO TIMBRE STUDY OF FORCED VOCALICS

YOSHIYUKI OCHIAI and TERUO FUKUMURA

Department of Electrical Engineering

(Received October 31, 1957)

Introduction

Among the problems touching upon the nature of the timbre of speech sounds, one of the most difficult to study and yet one which demands attention is found in the field where the relation between what are known as "quality" and "quantity" is to be debated. This relation is most discernible when speech sounds undergo a change in the uttering parameters of forcing volume (forcing level) and voicing pitch. As to timbre pattern of the sound produced, an increase of uttering force does not mean uniform reinforcement such as given by, for example, a distortionless amplifier, and a change in uttering pitch does not mean a simple uniform shift in frequency dimension as given by, for example, "Revolutional Synchronous Distortion" (RSD). Experiences with telecommunication teach us that even simple uniform amplification and transition can be a difficult study when we treat them from the viewpoint of timbre quality. Much more difficult may be the study of timbre changes caused by uttering conditions, because changes in timbre construction under uttering conditions are usually far more complex. But we face the real difficulty when the full significance of the aspect of dual-quality is demanded. We can even state positively that the difficulty and significance of our study arise when we take into account dual quality. Dual aspect is important as the very basis of our quality theory. We give a brief description of dual aspect because of its importance as an introduction to the subject of forced vocalics and because this aspect is important to us in our search among the branches of science and art and in the fields of engineering which are brought into contact through the subject of forced vocalics.

Consideration from Phonetic Aspect

Emphasis and de-emphasis of speech expression in conversation are most natural phenomena. In conversation, expression has several phases such as the one touching upon phonetic usage and colloquial custom, and the so-called rhetorical and aesthetic one, all related to the art of speaking. When we turn our attention to the expression in a more advanced way, we must look for aid in the art of singing. But when we limit our scope by restricting ourselves to phonetics in normal conversation, we can find the subjects of accent and intonation as bearing on our present subject of forced vocalics. As to accent, there are three kinds: Pitch, stress and elongation.¹⁾ Accent by lengthening the duration of sound is important in some languages, for example, in Russian and old Greek; but it is of secondary importance in the majority of languages where pitch and stress are of primary importance.

Technological and Psychological Considerations

Now we approach this forcing subject from another standpoint, *viz.*, that of speech telecommunication. Daily experiences in ordinary communication remind us of the fact: When the speaker is situated in surroundings which force him to deal with the inconvenience of interferences such as ambient noise and cross-talking, or when he is forced to talk over a telephone line with a poor connection (consequently causing an excessive heavy attenuation of speech), he must either increase his voice volume or raise his voicing pitch or, as the majority of cases may be, raise both pitch and volume at the same time to overcome the difficulty of his situation, or more exactly expressed, to react or adapt himself to his situation. In the specialized consideration of communication engineering, what is meant by these changes in the subjective condition of utterance which are shown as the reactions of the speaker to the objective surroundings of his telephoning? By his forcing the level of utterance, the so-called signal-to-noise ratio would be improved, and the effort made to raise the pitch would result in lessening (or mitigating) the so-called masking effect of noise disturbance. These technical answers seem reasonable, but they are somewhat formulary and a little superficial. From the psychological standpoint of speech quality, more pertinent and coming-to-the-quick answers must be given. It is only from quality theory, particularly from the quality theory of timbre aspect that the most satisfactory interpretation can be expected.

In the natural course of events, the subject of vocal forcing on uttering level touches upon the aspect of loudness as its primary and essential feature; the study of loudness in vocal forcing is neither an uninteresting nor a difficult one. We are, however, postponing this study of vocal forcing in loudness aspect until some future time. It is not in loudness but in timbre aspect that we now wish to detect influences of vocal forcing. This side of our study is without doubt far more difficult yet more important for illuminating the quality theory of dual aspect than any other. Vocal pitching should, of course, be included in a consideration of the subject of vocal forcing because pitch and stress are inseparable elements of intonation, as a rule.

Dual Aspect of Forced Vocalics—Semantic and Aesthetic

From what has been described in the foregoing, we are naturally led to the finding of dual aspect. That is, we arrive at a dual meaning of our subject of forced vocalics in speech expression. Forced vocalics bear dual meaning, *viz.*, the semantic and the aesthetic.²⁾ We can easily find the semantic meaning from the viewpoint of phonetics. On the other hand, in the transference of emotional information as found in expressional emphasis, for example, in dramatic art, we are able to find the aesthetic meaning. When we concern ourselves with emphasis of expression, we should perhaps allude to various effects of vocal forcing in the art of singing where in addition to the ordinary emphasis and de-emphasis expressed by *crescendo* and *diminuendo*, the most supreme technique on timbre nuance does exist.

In short, the significance of the semantic meaning of forced vocalics is conspicuous and considerable but it should not be overrated by ignoring the aesthetic meaning of forced vocalics which portrays emotional information. By these paths we are led unerringly to our home ground of quality theory. After all, it would

seem that the wisest plan would be to approach this subject of forced vocalics from the timbre aspect of dual quality, with which we have already familiarized ourselves.

Consideration from Timbre Aspect of Dual Quality

As a natural outcome of our intensive investigation on the general nature of the subject of forced vocalics, we came face-to-face with the dual aspect of quality. This means: In order to consider the subject in its broadest sense, that is to say, by acknowledging that forced vocalics have some bearing on human speech expression in general, not confining ourselves to one-sided intellectual information only, but extending our outlook further to include emotional information as well, it would be wise to treat the subject from the standpoint of timbre quality of dual aspect.

We had already shown in our previous papers³⁾⁴⁾ that phonemic and vocal patterns are best presented as a display of timbre. In these papers, the phonemic pattern is related closely to articulation quality, and vocal pattern is intimately connected with naturalness quality. Further, in a more advanced and definitive way, we introduced what we term "phonemic formant" as the characteristic and effective center of phonemic quality in phonemic pattern. At the same time we introduced "vocal formant" as the characteristic part of the structure of vocal pattern.

In defining such phonemic and vocal elements, we seriously considered what relation exists between the timbre-pattern display and the forcing phenomena of vocalics. In our process of obtaining patterns, the effect of pitch forcing was taken into full account. Phonemic patterns in general are subject to the least influence of pitch forcing and phonemic formants particularly are independent of pitch forcing insofar as the pitch range in natural conversation is concerned. According to our theory, on the contrary, vocal pattern must depend on pitch as a reasonable conclusion. In reality, it appeared to us that we must check up on the effect of volume forcing in utterances concerning the process of our pattern determination.

As this may be, the phenomena of forcing exhibit nothing astonishing when the forcing remains within the pitch compass and the dynamic range, as occurs, for example, in natural conversation. Even though there may be some or no forcing influence upon timbre qualities, there is no surprising effect. The most interesting manifestation is when the forcing surpasses the natural range and compass of normal conversation and extends itself into the range and compass of vocal music. For example, when pitch range extends over three or more octaves, the effect of pitch upon phonemic pattern becomes serious because it is possible that the pitch itself may pass through the so-called first formant region and go beyond. This means that the excessively highpitched voice may lose its phonemic value and there may be the same effect also when the voice is either too loud or too soft. In vocal music, however, we may consider that what of phonemic value is lost in semantic meaning is quite compensated for by the gains in aesthetic meaning.

If we are not mistaken, B. Stout⁵⁾ was the first to study forcing effect from the standpoint of vocal music. After him, T. Yamashita and M. Tsuda⁶⁾ of our Laboratory carried on the same study without trying to develop timbre study of forced vocalics. That was left for later research.

Schedule of Research

From past considerations and discussions, we are able to understand the kind of difficulties with which our present subject confronts us. We must naturally deal with fine structure in timbre pattern when studying timbre details of forced vocalics. This was not easy because to look at fine structure of timbre from dual aspect, that is to say, to carefully examine the phonemic and vocal patterns simultaneously, really imposed lengthy and time-consuming work upon us. To overcome this handicap and to reach our goal, necessitated some deliberately planned suitable and elaborate experiments. First we restricted our objective to the study of only one caller's voice, in exactly the same manner as in our nasal study. In our analyses, it seemed most fit to use the buccal egression of vocalics in acoustic oscillation waves because these oscillatory waves are the very substance of transmission in speech communication. For the purpose of clarifying the various influences of forcing upon fine structure of timbre and for the further purpose of segregating the separate influences on phonemic and vocal patterns, we are wise and prudent to adopt another advanced process, the vibration-detecting device of throat pick-up. From the nature of things, however, we should not overrate the merit of the physical aspect of timbre analysis.

Importance of Subjective Study

As already stressed in nasal study,⁷⁾ we must repeat the importance of subjective study on the problem of forced vocalics. Rather we should say that subjective study in forced vocalics is of far greater importance than in nasalics. This is because, from the timbre aspect of dual quality (semantic and aesthetic), the phenomena of forcing must be interpreted as relating to both semantic and aesthetic expressions; that is to say, to phonemic and vocal information. By using our terms, the subject of forcing must be studied in connection with phonemic and vocal patterns. To verify that phonemic pattern in general, phonemic formant in particular, and vocal pattern in general, and vocal formant in particular, are all intimately and significantly related to forcing phenomena, we must resort to the unerring verification furnished by subjective quality measurement. In other words, to determine whether the forced vocalics can have a bearing on vowel (phoneme) articulation or voice (vocal) naturalness, depends entirely on the results of subjective measurement.

As usual in our case, we advance subjective study of quality measurement parallel to objective study of timbre-pattern representation. To justify our interpretation of subjective data, use of this process is really indispensable. For this purpose, it is most convenient to adopt the type of distortion called band-eliminating distortion (BED) because by introducing distortion BED, we can clarify the inner mechanism of quality formation of speech in frequency dimension.

Does volume forcing bring about a gain or cause a loss in either articulation or in naturalness? This question is the solid form of the forcing problem. We can also pose the question: Does pitch forcing result in gain or loss in either articulation or in naturalness? Subjective experiments for the purpose of finding answers to these questions can give us not only reliable answers to the questions posed in our section "Technological and Psychological Considerations" but can also illuminate for us the essential nature of human information,

Postscript

Coincident with the start of our nasal study in 1953, we carried out our original experiments on forced vocalics to elucidate our dual-quality theory and to explain our definitions of phonemic and vocal elements. As to the experimental results, in 1955 we published one report⁸⁾ on timbre change as viewed from the standpoint of phonemic element and in the same year we published another report⁹⁾ viewed from vocal standpoint. At that time we could not accomplish a complete interpretation of the timbre change which is due to forcing. We recently came across two studies touching upon this subject, one¹⁰⁾ from the United States and the other¹¹⁾ from Russia, both of which were published in 1956. In these two studies, the subject was treated from the standpoint of articulation only. Under these circumstances, we feel it our duty to publish our detailed study in the form of a series report. In spite of the fact that we are still far from a perfect understanding of the phenomena of forced vocalics, we think we should make these reports available because of their broad coverage of timbre aspect.

Acknowledgement

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

References

- 1) M. Durand: *Voyelles Longues et Voyelles Brèves — Essai sur la Nature de la Quantité Vocalique*. LIBRAIRIE C. KLINCKSIECK, PARIS, 1946.
- 2) Y. Ochiai: *Mémoire sur les Sons des Voix Humaines*. MFE of Nagoya University, Vol. 4, No. 1, July, 1952.
- 3) Y. Ochiai and T. Fukumura: *Timbre Study of Vocalic Voices*. MFE of Nagoya University, Vol. 5, No. 2, Sept., 1953.
- 4) Y. Ochiai und T. Fukumura: *Beiträge zur Erkenntnis der Klangfarbestruktur bei vokalischen Klangbildern*. MFE of Nagoya University, Vol. 8, No. 1, May, 1956.
- 5) B. Stout: *The Harmonic Structure of Vowel in Singing in Relation to Pitch and Intensity*. JASA, 10, 2, 1953.
- 6) T. Yamashita and M. Tsuda: *Some Observations on Voicing Area in Vocalization Phenomena*. MFE of Nagoya University, Vol. 6, No. 1, May, 1954.
- 7) Y. Ochiai: *Memoirs on Nasals*. MFE of Nagoya University, Vol. 9, No. 1, May, 1957.
- 8) Y. Ochiai and T. Fukumura: *Hassei-Kyodo ni yoru Onsei-Pattern no Henka ni tsuite (Japanese)—On Timbre Change Caused by Uttering Force*. Read at Joint Meeting of Societies of Elect. Eng., Comm. Eng. and Illum. Eng. of Japan, April, 1955.
- 9) Y. Ochiai and T. Fukumura: *Hassei-Kyodo ni yoru Boin-Onshoku-Kozo no Henka ni tsuite (Japanese)—On Changes in Timbre Pattern of Vocalics Due to Uttering Force*. Read at Meeting of Acoustic Society of Japan, May, 1955.
- 10) J. M. Pickett: *Effect of Vocal Force on Intelligibility of Speech Sounds*. JASA, 28, No. 2, 1956.
- 11) М. А. Самсонов: *Спектр Русской Речи и Его Деформация под Воздействием Сильного Звукового Раздражителя. Восприятие Звуковых Сигналов в Различных Акустических Условиях*. ИЗДАТЕЛЬСТВО АКАДЕМИИ НАУК СССР 1956.