

TRANSMISSION OF QUALITY

YOSHIYUKI OCHIAI

Department of Electrical Engineering

(Received October 13, 1954)

Guided by a practical study of the so-called transmission quality, we have discussed on the conception of *quality*, the most fundamental and proper conception in communication engineering, making a fine contrast with the conception of *efficiency* that prevailed so supremely in whatsoever science of engineering. Focussing many a diverged step of thinking, we have arrived, in the end, at the standpoint of "subjective phonal aspect," of which the outline sketch has been tried under the motive idea that quality transmission must be differentiated from wave-form transmission.

Introduction

It might be difficult to determine the exact date of birth of the conception of *quality* in the history of telecommunication. But the possible inference is as follows: the conception of quality did not appear until the development of the engineering of telecommunication could dispose of the attenuation-distortion, that is, the most disturbing distortion in the prehistory of communication engineering. The appearance of the quality-conception is of great importance for communication. It is to be noticed. It means that the conception of *efficiency*, the most supreme conception in the sciences of engineering in general has to be replaced by the conception of quality so far as the science of communication is concerned. Such an alternation of the fundamental conceptions is of hot interest. According to what history tells us, the discovery of conception of the quality was made only with much difficulty. But one is apt to forget what quality really is and is for in communication. It is therefore needful to try a conceptual study at the beginning. A misconception must be removed from the path.

Quality in Communication

We must discuss here most fundamentally the conception of quality in communication, keeping chiefly in mind of the transmission problem in speech communication. The conceptual study of quality is nothing but an interpretation on quality that is necessarily connected with an interpretation on communication itself. After so many turns and twists of consideration, we arrived at the stage of thinking which is: the quality is a conception that reflects directly the essence of communication; this conception cannot be derived without considering the subjective circumstances constituted with human expression and impression as a talker and listener in communication, an indispensable part in communication phenomena. The conception of quality in physical sense, that is, quality definable by pure physical processes is in itself but nonsense. The true meaning or the full savour of quality does extinguish wherever and whenever the subjective side of human factors

comes to be excluded. We call such an aspect "subjective phonal aspect."

Description by Symbolic Method

For proceeding with our quality theory in subjective phonal aspect, it will be fit to adopt some convenient method such as a symbolic method. It is because the consideration and description of the so-called quality phenomena that is the most complicated problem, containing both an objective side and a subjective side of things, will become very difficult in a usual expression or explanation in words. The description by words does not really carry us very far. For the sake of economy of words in description, and moreover for avoiding the danger to miss the trail of logics in the labyrinth of treatment of so complicated problem, the adoption of an operational symbolical method will be quite pertinent and serviceable.

Fundamental Expression

The present purpose of our quality study in telecommunication is to make clear the nature of speech sounds as a sort of transmission effect caused by transmission processes. In other words, it must be our purpose to clarify the functional character of the transmission systems, basing upon the various phases of the distortion-phenomena of the speech signal impressed upon them and reproduced by them. Namely the direct aim of our study is nothing but an observational measurement on the speech sounds disturbed and distorted by the transmission systems, that is, the study of the *transmission effect* of transmission systems by means of an investigation of the degree of degradation of speech quality of which the phenomena can be recapitulated by name of *quality response* of transmission-systems. We cannot say distinctly how the effect of transmission systems really is until we can assert the degree of distortion that is observable only in the feature of speech sounds transmitted. We must admit at any rate that there exists, in fact, an indivisible relation between the transmissional means and the transmittable things. We ought, therefore, to have an expression on the fundamental function of speech transmission-systems for the speech signal impressed, transmitted, and reproduced.

It will be suitable en passant to make a notice on the speech signals themselves. We are inclined to take the actual complex speech sounds as signal; we are not in a position to take as signal the so-called pure tone of sinusoidal nature that used to be employed as such in testing, for example, the frequency response characteristics of the systems. It goes without saying that we cannot also overlook the rôle of the human sensation in the qualificative interpretation of communication phenomena. To cut matters short; by what means one comes to know whether or not the disturbing effect of transmission process really exists? Without appealing to the human sensation, can we say anything reliable about the so-called *transmission quality*? Here ought to come into question the psycho-physiological factor as a kind of mental and sensational judgement on the transmission effect of the transmission systems. In amount, we must bind all these factors together in a unitary viewpoint in speech communication. We will try a symbolic description in consideration on quality theory, of which the most fundamental expression is firstly to be set down. Denote by **D** the physical operation as to the disturbing distortion of transmission-system, and by **H** the hearing or judging opera-

tion. Denote by X the input signal in physical sense, and by \mathcal{E} the output signal in subjective sense. Employing these notations, we suppose to show the fundamental definition on *transmission quality* or *system quality* by

$$\mathbf{H} \cdot \mathbf{D}(X) = \mathcal{E}',$$

and also to define the *speech quality* by the expression

$$\mathbf{H}(X) = \mathcal{E}.$$

By adopting such fundamental expressions, we can proceed to a general form of measure on quality transmission in subjective sense by the following expression

$$\Gamma = \mathbf{H} \cdot \mathbf{D}(X) : \mathbf{H}(X),$$

where the sign: means the procedure of some comparison of subjective nature. If we are forced to lead the quality expression from the point of view where the subjective considerations are perfectly excluded, then we shall have

$$G = \mathbf{D}(X) : (X)$$

where the sign: implies the procedure of some objective comparison.

Phonal Qualities

Subjective appreciation of transmission effect must depend upon the judgement on qualities which correspond to what is called phonal attributes. It means that the phonal attributes must be chosen as a foundation on which all sorts of subjective criticism are carried out; we have to take the subjective ground such as phonal attribute in our quality interpretation, instead of taking the physical ground of wave form as in the wave-form transmission theory. It must be justly on this ground that transmission phenomena are examined and the transmission effects are appreciated, and transmission systems or processes are really rated. This is the idea of quality-transmission. Guided by such a thinking, we can and must have three qualities: loudness quality, pitch quality and timbre quality. These are phonal qualities which we can make use of in our theory. We will explain by quoting an actual example. In the subjective expression of speech quality

$$\mathbf{H}(X) = \mathcal{E},$$

if the speech signal X is supposed to be subject to the mental judgement that depends upon the three phonal attributes, then we can give the forms of three qualities of speech signal:

$$\mathbf{H}_p(X) = \pi(\mathcal{E})$$

$$\mathbf{H}_l(X) = \lambda(\mathcal{E})$$

$$\mathbf{H}_t(X) = \tau(\mathcal{E})$$

where the suffixes p , l , t of the operation \mathbf{H} signify that the attitude of recipient depends mainly on the pitch-, level- and timbre-element respectively, and $\pi(\mathcal{E})$, $\lambda(\mathcal{E})$ and $\tau(\mathcal{E})$ mean respectively subjective pitch, subjective loudness, and sub-

jective timbre of subjective signal \mathcal{E} . In appreciation of the point of judgement, we can call these quality-judgements respectively p -judgement, l -judgement and t -judgement.

Quality Transmission

If we admit that the quality transmission, as we have shown, is expressed in general by

$$\Gamma = \mathbf{H} \cdot \mathbf{D}(X) : \mathbf{H}(X),$$

then we can give more precisely and concretely the forms of individual quality. For instance, as to the loudness quality, we have the form of loudness transmission:

$$\Gamma(\lambda) = \mathbf{H}_l \cdot \mathbf{D}(X) : \mathbf{H}_l(X),$$

and for pitch transmission, we have

$$\Gamma(\pi) = \mathbf{H}_p \cdot \mathbf{D}(X) : \mathbf{H}_p(X),$$

and lastly, for timbre transmission, we have

$$\Gamma(\tau) = \mathbf{H}_t \cdot \mathbf{D}(X) : \mathbf{H}_t(X).$$

There are so many types of distortion. We can consider not only the level-distortion but the pitch- and timbre- distortion. It is not in vain, therefore, to consider the three individual cases of quality transmission.

Quality Comparison and Quality Matching

For developing the conception of quality transmission with a view to the practical application, it is convenient to guide the derived ideas of *quality comparison* and further *quality matching*. These two conceptions are alike; they do not differ so much. But it looks needful to differentiate one from another from the viewpoint of practical measurement. In our notation of quality transmission

$$\mathbf{H} \cdot \mathbf{D}(X) : \mathbf{H}(X)$$

we have said that the sign: means some comparative procedure of subjective nature. This is but the idea of quality comparison. By quality comparison we understand, for example, the greater-or-smaller judgement in loudness or the higher-or-lower judgement in pitch. So far as the content or material of quality itself is susceptible of measuring and scaling, the comparative judgement on quality is naturally possible. Even if we are not successful in scaling or grading the magnitude of quality, we can do nevertheless execute quality comparison, so long as a sort of order such as greater-or-smaller or higher-or-lower can be established as to the quality concerned. On the contrary, the conception of quality matching is not established until the process of scaling for the magnitude of quality can be prepared. Suppose that we can find some convenient reference system easily calibrated and continuously and smoothly adjustable from zero to very large magnitude of distortion, which permits us to measure in scaling any distortion of interest. By utilising such reference \mathbf{D}_r , we can carry out the quality matching of the

system D , as the next relation shows

$$H \cdot D(X) : H \cdot D_r(X).$$

We will illustrate these examples more concretely regarding the individual quality. By aid of the conception and technique of loudness matching of systems, we can attain the conception and the practical measuring technique concerning the loudness matching of speech sounds themselves. According to the notation prescribed, the loudness comparison of any two sounds X_1 and X_2 is expressed by

$$H_l(X_1) : H_l(X_2).$$

By utilising the reference system D_r , and satisfying the following relations

$$H_l(X_1) = H_l \cdot D_{r(1)}(X'_1)$$

$$H_l(X_2) = H_l \cdot D_{r(2)}(X'_2)$$

$$X'_1 = D_0(X_1)$$

$$X'_2 = D_0(X_2)$$

wherein D_0 implies certain distortion necessary for approaching the matching process, we can lead the loudness comparison of sounds to the direction of loudness matching.

Pitch Matching between Systems and between Phones

The pitch comparison of two systems D_1 , D_2 is expressed by

$$H_p \cdot D_1(X) : H_p \cdot D_2(X).$$

If we succeed in finding some convenient reference system D_r for reference of pitch comparison so as to satisfy the following relations

$$H_p \cdot D_1(X) = H_p \cdot D_{r(1)}(X),$$

$$H_p \cdot D_2(X) = H_p \cdot D_{r(2)}(X),$$

then we are in a position to arrive at the conception of pitch matching of the transmission systems by the following expression

$$H_p \cdot D_1(X) : H_p \cdot D_2(X) = H_p \cdot D_{r(1)}(X) : H_p \cdot D_{r(2)}(X).$$

Furthermore, by means of reference system D_r and accessory system D_0 , we can proceed to the conception of pitch matching of sounds X_1 , X_2 by the expressions

$$H_p(X_1) : H_p(X_2) = H_p \cdot D_{r(1)}(X'_1) : H_p \cdot D_{r(2)}(X'_2),$$

and

$$X'_1 = D_0(X_1), \quad X'_2 = D_0(X_2).$$

Loudness Comparison and Loudness Matching

Depending on l -judgement, the conception of loudness comparison of any two distortion systems is expressed in general by the form

$$\Gamma_1(\lambda) : \Gamma_2(\lambda) = H_l \cdot D_1(X) : H_l \cdot D_2(X).$$

When we succeed in finding some sort of distortion D_r that is suitable for reference of loudness comparison, we can obtain the following relations as to any series of distortion $D_1, D_2, D_3, \dots, D_n$

$$\begin{aligned} H_l \cdot D_1(X) &= H_l \cdot D_{r(1)}(X), \\ H_l \cdot D_2(X) &= H_l \cdot D_{r(2)}(X), \\ H_l \cdot D_3(X) &= H_l \cdot D_{r(3)}(X), \\ &\dots\dots\dots, \\ H_l \cdot D_n(X) &= H_l \cdot D_{r(n)}(X). \end{aligned}$$

In such a manner we can settle the conception of loudness matching. Loudness-transmission effect of transmission systems D_1, D_2, D_3, \dots is precisely evaluated by numerical data of reference system D_r , such as $D_{r(1)}, D_{r(2)}, D_{r(3)}, \dots$

General Consideration on Timbre

Let us now turn to the last quality, infinitely more complicated nature, that is, the timbre quality. Timbre is the most important of all three qualities. Timbre is the most difficult to be understood. It seems to be a sort of quality that is concerned with a conception of *a whole*. Nevertheless we must make much of timbre in the consideration of quality transmission. Compared with the timbre quality, the loudness quality is of little importance. As to the pitch quality, the pitch study might reasonably be made the indispensable prelude to the subtle study of timbre problem. Pitch is an important quality next to timbre. Timbre is matchless in this point. The importance of timbre quality cannot be valued too highly.

From the viewpoint of subjective aspect, we must consider the timbre in the subjective meaning. In such a manner of thinking, the conception of timbre quality becomes too vast and too complex. It is true. But we cannot neglect the timbre study because of its complexity, nor give up the subjective aspect because of its vastness.

There are so many problems primary and essential on timbre quality: for instance, the problem of its definition, and the problem of attributal localisation, *i.e.*, the problem of determining the relationship between attributes, and the problem of measuring the timbre pattern. We have no intention to deal with all these problems at a time. We shall have another chance to report the more detailed study on timbre. Here we must treat the timbre problem so as to conform to the frame of quality transmission. We must primarily treat the problem of timbre comparison, and if it is possible, further proceed to the most difficult problem of timbre matching. Before entering upon these studies, we will consider the phenomena of pure discrimination as an introductory and preliminary study.

Differential Sensation of Pure Tone

By symbolic method, we will try here to note the functional phenomena of differential sensation, and will try, if necessary, to give fine distinction between various conceptions thereof. Note the pure tone of sinusoidal nature with $x \binom{i}{f}$ where i means its intensity, f its frequency, both in physical expression. Then

we have the followings as quality expressions of pure tone

$$H_l \left[x \left(\begin{matrix} i \\ f \end{matrix} \right) \right] = \lambda(\xi),$$

$$H_p \left[x \left(\begin{matrix} i \\ f \end{matrix} \right) \right] = \pi(\xi).$$

For giving the process of differential sensation, we can utilise the form of so-called quality comparison. For intensity discrimination, we have

$$H_l \left[x \left(\begin{matrix} i \\ f \end{matrix} \right) \right] : H_l \left[x \left(\begin{matrix} i \pm \Delta i \\ f \end{matrix} \right) \right] = [\lambda]_f : [\lambda + \Delta \lambda]_f,$$

and for pitch discrimination,

$$H_p \left[x \left(\begin{matrix} i \\ f \end{matrix} \right) \right] : H_p \left[x \left(\begin{matrix} i \\ f \pm \Delta f \end{matrix} \right) \right] = [\pi]_i : [\pi \pm \Delta \pi]_i.$$

It comes usually into question to determine the value of $\Delta i/i$ or $\Delta f/f$ that corresponds to the discernible increment of sensation. Usual process is to find the value of $[\Delta i/i]$ corresponding to the subjectively discernible increment of loudness $[\Delta \lambda]_{f=\text{const}}$, and also the value of $[\Delta f/f]$ corresponding to $[\Delta \pi]_{i=\text{const}}$. It means here that the two sounds for loudness comparison must be matched in frequency, and that the two sounds for pitch comparison must be matched in intensity. It is but the stage of physical matching that is prepared for subjective discrimination. Speaking more strictly, we ought to have $[\Delta i/i]$ that corresponds to $[\Delta \lambda]_{\pi=\text{const}}$, and to have $[\Delta f/f]$ that corresponds to $[\Delta \pi]_{\lambda=\text{const}}$. Here it is meant: the two sounds for loudness comparison must be matched in pitch, and the two sounds for pitch comparison must be matched in loudness. Here is to be prepared the process of subjective matching for subjective discrimination. We have thus two different methods: one is the subjective discrimination under physical matching, and the other is the subjective discrimination under subjective matching. Principally the former differs a little from the latter. We should like to call the former case the *almost pure discrimination*, and to call the latter case the *pure discrimination*. Lastly, it is to be added that there can be the case of discrimination where the differential sensation of one element is obtained without any matching process of other element. We must call this last case the *discrimination in general*.

Timbre Comparison

Of so many studies on quality transmission, the most important and difficult one is but the study of timbre comparison. The timbre comparison of two systems or two sounds will be denoted respectively by

$$H_t \cdot D_1(X) : H_t \cdot D_2(X) = \tau' : \tau'',$$

or

$$H_t(X_1) : H_t(X_2) = \tau_1 : \tau_2.$$

In order to study more closely the problem of timbre comparison, we are forced to think of the treatment of other qualities than timbre. But, before going on to

the detailed considerations, it is very fit to think a little about the content of t -judgement on which the process of timbre comparison quite depends. We can never expect therein a judgement of greater-or-smaller type. In the process of timbre comparison we see a judgement of similar-or-unsimilar type works. It is rather beside the question, whether or not in the timbre comparison the brighter-or-darker, or thicker-or-thinner, or heavier-or-lighter judgement can be settled. In any case, the greater-or-smaller judgement is quantitative; and the similar-or-unsimilar judgement is rather qualificative. We will try a simple description on timbre comparison between two sounds without much rigorousness. In timbre study we shall adopt the following notation for the sake of convenience

$$\mathbf{H}_t \left[X \begin{pmatrix} l \\ p \\ t \end{pmatrix} \right] = \mathbf{H} \left[T \begin{pmatrix} l \\ p \end{pmatrix} \right] = \tau \left(\frac{\lambda}{\pi} \right) = \tau(\Xi)$$

where T means speech signal in timbre expression. Therefore, instead of the expression

$$\mathbf{H}_t \left[X_1 \begin{pmatrix} l_1 \\ p_1 \\ t_1 \end{pmatrix} \right] : \mathbf{H}_t \left[X_2 \begin{pmatrix} l_2 \\ p_2 \\ t_2 \end{pmatrix} \right],$$

we have, as a convenient form of timbre comparison

$$\mathbf{H} \left[T_1 \begin{pmatrix} l_1 \\ p_1 \end{pmatrix} \right] : \mathbf{H} \left[T_2 \begin{pmatrix} l_2 \\ p_2 \end{pmatrix} \right] = \tau_1 \left(\frac{\lambda_1}{\pi_1} \right) : \tau_2 \left(\frac{\lambda_2}{\pi_2} \right).$$

when λ_1 is not equal to λ_2 and π_1 is not equal to π_2 , we shall have timbre comparison in general sense, that is, the case of $[\tau_1/\tau_2]_{\substack{\lambda \neq \text{const} \\ \pi \neq \text{const}}}$, which means timbre comparison under the condition of non-matching in loudness and pitch.

Now, by aid of some convenient processes toward T_2 , we can assume that the following relations are satisfied

$$\tau_2' \left(\frac{\lambda_1}{\pi_2} \right) = \mathbf{H}_t \cdot \mathbf{D}_{r(l)} \left[X_2 \begin{pmatrix} l_2 \\ p_2 \\ t_2 \end{pmatrix} \right] = \mathbf{H} \left[T_2' \begin{pmatrix} l_1' \\ p_2 \end{pmatrix} \right],$$

and

$$\tau_2'' \left(\frac{\lambda_1}{\pi_1} \right) = \mathbf{H}_t \cdot \mathbf{D}_{r(p)} \cdot \mathbf{D}_{r(l)} \left[X_2 \begin{pmatrix} l_2 \\ p_2 \\ t_2 \end{pmatrix} \right] = \mathbf{H} \left[T_2'' \begin{pmatrix} l_1' \\ p_1' \end{pmatrix} \right],$$

then we have, as timbre comparison in pure sense

$$\tau_1 \left(\frac{\lambda_1}{\pi_1} \right) : \tau_2'' \left(\frac{\lambda_1}{\pi_1} \right) = \mathbf{H} \left[T_1 \begin{pmatrix} l_1 \\ p_1 \end{pmatrix} \right] : \mathbf{H} \left[T_2'' \begin{pmatrix} l_1' \\ p_1' \end{pmatrix} \right],$$

which means timbre comparison under the condition well-matched in loudness and pitch, that is, the case of $[\tau_1/\tau_2'']_{\substack{\lambda = \text{const} \\ \pi = \text{const}}}$. When so much strictness is not needed,

we can have the case of $[\tau_1/\tau_2']_{\substack{l = \text{const} \\ p = \text{const}}}$, that is,

$$\tau_1 \left(\begin{matrix} \lambda_1 \\ \pi_1 \end{matrix} \right) : \tau_2'' \left(\begin{matrix} \lambda_1' \\ \pi_1' \end{matrix} \right) = \mathbf{H} \left[T_1 \left(\begin{matrix} l_1 \\ p_1 \end{matrix} \right) \right] : \mathbf{H} \left[T_2'' \left(\begin{matrix} l_1 \\ p_1 \end{matrix} \right) \right]$$

where

$$\tau_2'' \left(\begin{matrix} \lambda_1' \\ \pi_1' \end{matrix} \right) = \mathbf{H}_t \cdot \mathbf{D}_{r(p)} \cdot \mathbf{D}_{r(l)} \left[X_z \left(\begin{matrix} l_2 \\ p_2 \end{matrix} \right) \right] = \mathbf{H} \left[T_2'' \left(\begin{matrix} l_1 \\ p_1 \end{matrix} \right) \right].$$

This case corresponds to timbre comparison in almost pure sense. In the above, the notations $\mathbf{D}_{r(p)}$ and $\mathbf{D}_{r(l)}$ imply respectively reference system for pitch matching and reference system for level matching. For avoiding the complexity of description, the explanation was simplified by neglecting the unimportant process for comparison. It needs scarcely be said: when the pitch matching is carried out on the side of uttering subject, the preliminary process for matching becomes only one, that is, the level matching.

Timbre Matching

The development of timbre theory depends entirely on the practical technique of timbre matching. The problem of timbre matching is so difficult that it has never been considered, and we can not also answer it with ease. We think thus: as the conception of timbre is so difficult that we cannot find the real technique of timbre matching. But we can say also in such a manner: for no other reason than that the technique of timbre matching is not yet discovered, the conception of timbre cannot yet be distinctly established. On turning our attention to the actual problem of transmission rating in telecommunication, especially in speech communication, which is reduced, after all, to the problem of timbre appreciation in transmission, we see so many treatments without fundamental knowledge on timbre and also the basic technique of timbre matching. This is just to put the cart before the horse. In amount, the problem of timbre matching is a pending question to be solved.

Discussion

Whenever we meet the practical measurement of articulation in speech communication, we always are induced to conceive: it must be a systematical thinking on quality as a whole, not crumbs of its information as occasions demand, that we really need in studying such a problem as quality. By faithfully conforming to the traditional, conventional quality principle, we shall not be able to find a procedure of quality measurements on naturalness, another sort of quality than articulation, even though we shall be proficient in practical measurements in articulation quality: if we follow the traditional steps by treating the pitch as quite different from the loudness, we shall be unable to invent the precise measurement as to the subjective pitch of complex sound signal, even though we shall become skilful in measurement in loudness; we must be afraid above all that the problem of transmission quality (for example, the articulation in speech communication) can practically be disposed of without touching upon the problem of timbre. It must be a misconception that we imagine to achieve a qualificative measurement on timbre without obtaining the fundamental knowledge on timbre. We treat here three qualities, pitch, loudness, timbre, preparing the common basis of unitary

aspect. This is the idea of quality transmission.

Conclusion

This is some systematic report attempted on quality study in speech communication viewed from the "subjective phonal aspect." According to the history of telecommunication the traditional idea of *transmission efficiency* has, for the most part, given way to a rising conception of *transmission quality* in accordance with the rapid development of transmission engineering. We can never overlook the importance of the rôle of subjective part in communication phenomena. It is precisely therefrom that the conception of quality is reasonably derived. Starting from such a standpoint and conforming to such a frame of aspect, we are obliged to solve the quality phenomena on transmission in connection with the subjective base of phonal attribute instead of the physical base of wave form.

As the detailed studies of quality transmission, we have endeavored to make clear the two conceptions: quality comparison and quality matching. We have studied further the problem of comparison and matching regarding pitch- and loudness-quality, selecting the practical case of transmission-system rating as an engineering problem on one hand and the case of speech-sound appreciation as a phonological problem on the other hand. As to the timbre quality, the problem of timbre comparison has been studied. We have distinguished three kinds of comparison: timbre comparison in general sense, timbre comparison in almost pure sense, timbre comparison in pure sense. The problem of timbre matching has only been proposed and suggested here.

What we have called here in this essay "speech sound" in speech communication, might more generally and properly be said "phonal signal" or "acoustical signal." For, in this abstract study the speciality of speech sound does not come into question.