RESEARCH REPORTS

GENERAL CONSIDERATION ON STUDIES OF SPEECH QUALITIES IN ROTATIONAL SYNCHRONOUS DISTORTION

Yoshiyuki OCHIAI

Department of Electrical Engineering

(Received May 13, 1955)

This is a study concerning speech quality by the *rotational* or *revolving synchronous distortion*, a sort of specific transmissional distortion coming from a state of departure from synchronism between sound-recording and -reproducing systems of tape-recorder. As to the distortion of this kind, we can do away with the pitch- and loudness-study, turning our interest exclusively to the problem of timbre transmission. All observations and measurements were executed, so to say, from a timbre aspect of speech quality. In answering first to "clearness" requirement in speech communication, the articulation characteristics in a general and collective sense were primarily studied. For subjective study on timbre-quality mishearing phenomena between phonemes were treated with intense interest. We made further investigations on the allowance of this distortion aiming at timbre-discrimination in the case of the slightest departure from synchronism. Finally, we carried on some naturalness measurement in this distortion from which we can lead interesting and instructive results serviceable for transmission engineering.

General Consideration

Nature of distortion

The distortion called here tentatively Rotational Synchronous Distortion (RSD in abbreviation) is essentially of *transitional* nature, i.e., a distortion which causes shift of position in total of the sound in the domain of frequency dimension. As for pitched timbre, this distortion brings about some pitch change but no slightest change on the internal structure of timbre. In this respect this distortion stands in contrast with the so-called Carrier Synchronous Distortion (CSD); the latter, of course, belongs to the one of transitional type but it causes the degradation of harmonicity of timbre. For timbre study on speech quality these two distortions (CSD and RSD) are both indispensable. Because, by comparing speech qualities caused by these two types of distortion, we can advance so far in the interpretation of timbre study. We already reported rather collectively on quality study by CSD. Now we have to finish our report of the study by RSD of which the experiment was already executed four years ago. Besides delicate effects on timbre quality, the distortion RSD is also marked by an effect upon the duration quality (*i.e.*,

time construction) of speech sounds which will be otherwise with the distortion CSD. The distortion RSD can make either long or short the duration time of speech signals according as it speeds up or down its revolution.

As for the distortion which causes only a pure position-shifting in frequency dimension without bringing any disturbances into the internal structure of sound, we must utilize it to the utmost for timbre study. We know as self-evident fact that the speech phones are to be so pronounced that they might be understood irrespective of the difference of spacial size of utterers' vocal organ. For example, a large mouth can pronounce so good vowels as a small mouth does. It means that the vowels hold good as such notwithstanding the absolute volume of mouth, and they seem to stand in a connection with the spacial proprotionality of the articulatory parts brought into play by the utterance of the vowels in question. But this may be a matter of relativity. Our actual ordinary experiences in phonological aspect give a series of facts which are brought to pass only in relatively narrow extent existing in reality. What becomes of that fact in an imagined case of so large extent that exceeds by far the reality? In answer to such a question the study of speech quality by the distortion RSD may come forth. Because, the up- or down-distortion of RSD can easily aid us in embodying any kind of *dwarf* or *giant* only imaginable but non-existent. Even to the most subtle problem of speech timbre which lies in a delicate relationship between the vocal-cord effect and the vocal-cavity performance, the speech study of this kind may be able to answer with something. But we cannot expect too much in this respect, for we have not yet enough knowledge as to the vocal- and phoneme-quality separately; much less to the problem of interdependence between them.

As we know, this distortion RSD is very often used in the field of broadcast engineering because of its speciality in acoustical effect. Our study, however, was executed, not only in the interests of practical application, but also from the scientific intention towards timbre problem. Timbre is considered a construction in frequency domain. This construction must be best studied only by the distortion as to the frequency element. Without an aid of this distortion, we cannot go on into a profound study of timbre.

Before setting to dwell on our study, let us take a glimpse on studies ever executed: In a book* published in 1929, H. Fletcher shows an articulation-characteristic for the distortion RSD given by the disc-recording equipment: The maximum articulation in normal state is no more than 88 per cent, and the speed change in rotational ratio is seen only in the range covering from about 0.5 as minimum to about 1.7 as maximum. V. Engelhardt and H. Gehrcke** published in the same year the paper titled "Über Vokale" where they describe the articulation characteristics of five German vowels and further touch upon some mishearing phenomena among vowels: The distortion range used is also from almost 1/2 to about 2 in rotational ratio. In our research, we performed a more systematic and more collective timbre study on speech-phone for the distortion of wider range than in the cases of Fletcher and Engelhardt. Firstly, we wanted a distortion that might give pitch-change of about six octave range. But various circumstances in

* H. Fletcher; Speech and Hearing p 294. Van Norstrand Co. 1929.

** V. Engelhardt und H. Gehrcke; Über Vokale, Zeitschrift für technische Physik. Nr. 11, s. 563, 1929.

machine construction prevented us from realizing such distortion range, permitting finally only the range from 1/3 to 8/3 in rotational ratio. We had 3 points in speed-up region and further 3 points in speed-down region. Thus obtaining 6 points of discontinuous changes in revolutional speed, and adding one point, i.e., a normal point, we had a total of 7 points ready to be tested.

Experimental plan

Our research should consist of four parts of experiment: the first, articulation test, the second, mishearing study, the third, allowance test, the fourth, naturalness measurement. Articulation test has been made, of course, from the standpoint of "clearness" aspect in speech communication; but the aim of our study in this branch not being restricted to the direct demand of the real field of communication, our observations and measurements have been done in the intense interests in the timbre problem of speech-phones, and accordingly we have endeavoured to contribute to an interpretation of subjective timbre in dealing with the "mishearing" phenomena in a relatively detailed manner.

Our attitude assumed in the study of allowance test is quite the same as before. The allowance test also came originally from the real demand of the broadcast engineering. But this contains very interesting items based upon the timbre discrimination, that is, the most inspiring problem in timbre study. Our attempt was to find some clues lying hidden in the discrimination of this kind.

As we have already shown, this distortion is profitable as an effective means for studying pitch and timbre changes. Consequently, we tried to make some closer investigation on this point.

Naturalness test was carried out in a view to make clear the quality concept of naturalness by aid of this distortion. In order to show clearly that the naturalness quality is quite other than articulation, it is very suitable to adopt this distortion by which we can see responses of two qualities (articulation and naturalness) appearing different evidently from each other. For such a proof this distortion is most serviceable.

Distortions of pure nature and compound nature

To detect the cleancut response of quality against the pure distortion of SRD, it is needful to check whether the distortion in question is really of pure nature or not. When the distortion RSD is accompanied by any other distortion, the quality responses obtained cannot boast of its purity. And, in our experiment, we count as other kinds of distortion that come to play: (1) Band eliminating distortion, (2) Non-linear distortion, (3) Wow-distortion caused by irregular revolutions, (4) Noise distortion of any kinds. As for the distortions (3) and (4), we can almost neglect them by careful installation and fine adjustment of the equipment. But the distortion (2) is more or less inevitable in the present stage of equipment and it cannot be perfectly removed with ease. All that we can and must do now is only to check whether it remains almost unvariable notwithstanding the change of the distortion in question. If it varies the magnitude of distortion in accompany with the change of distortion RSD, the quality response does not reflect the pure effect of RSD. The distortion (1) is also essential in our experiment; for, trans-

itional distortion is always and necessarily accompanied by band eliminating distortion. This is based utterly on the fact that the transmission characteristic of the equipment used is finite and restricted in frequency characteristic. For example, the transition in higher frequency-region means a cutting-down of the lower frequency characteristic (of the system concerned) just in proportion to its extention in higher frequency, and *vice versa*. We have taken the labor of characteristic-compensation for flatness only when extreme cases of the greatest transition are devised with some strains at the sacrifice of its flatness. For the precaution in such a situation, we have consciously undertaken further experiment in the distortion of compound nature where RSD and BED appear hand in hand. A Study of the distortion of compound nature will be available for the treatment of "Quality Analysis."