ARTICULATION STUDY OF SPEECH QUALITIES IN ROTATIONAL SYNCHRONOUS DISTORTION

YOSHIYUKI OCHIAI, SHUZO SAITO and YOSHITAKA SAKAI

Department of Electrical Engineering

(Received May 13, 1559)

It is brought into question; primarily articulation study in the distortion RSD of pure nature, and secondly similar study in a combined distortion of RSD and BED. These studies, of course, executed from the clearness aspect in speech communication, must be placed in a closer connection with the "phoneme" quality of speech. From a collective consideration based upon the articulation-characteristic study, we can lead some important conclusion about the question how speech-phones with or without pitch come to play with their clearness performances.

Experimental Procedure

Subjective system

Three persons were selected as talking subjects; one male (S.S., aged 29) whose voice has about 130~ pitch, one female (S.N., aged 27) whose voice-pitch is about 200~, one boy (F.S., aged 10) whose voice-pitch is about 210~. The talker was seated at the distance of about 50 cm from a microphone in the case of voiced utterance, and about 30 cm in whispered utterance. As listening subjects, three young male students with normal hearing were employed. Listeners were seated at the distance of about one meter from a loudspeaker in a sound-proof room.

Speech-phone system

As in the usual way were collected and tested one hundred syllable-unit sounds of CV construction which form an almost complete whole of the Japanese language phone system, containning in themselves 36 speech-phones ended with glided vowels and 64 phones ended with direct vowels. Filling one series of logatome with 200 phones (two sets of hundred phone system), we divided it into four divisions of 50 phones each; four cards make one series: and we prepared a total of six series of card for the articulation experiment. For one calling subject and one listener 1200 phones were used, and for one caller and three listners 3600 observations were executed to determine any one point of characteristics.

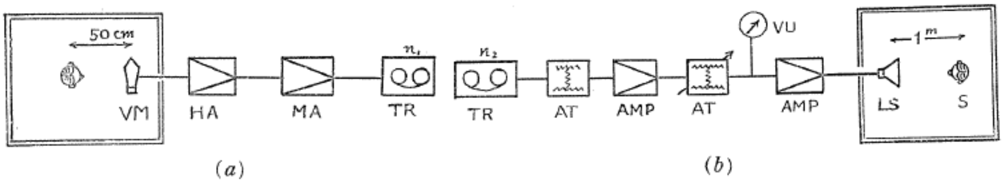


FIG. 1. Block-diagram of testing circuit systems employed for articulation measurement (*a*): recording system. (*b*): reproducing system.

Objective system

The recording and reproducing systems of tape-recorder equipment employed for the experiment are shown block-diagrammatically in Fig. 1, and the transmission characteristics of essential parts of the system are shown in Fig. 2.

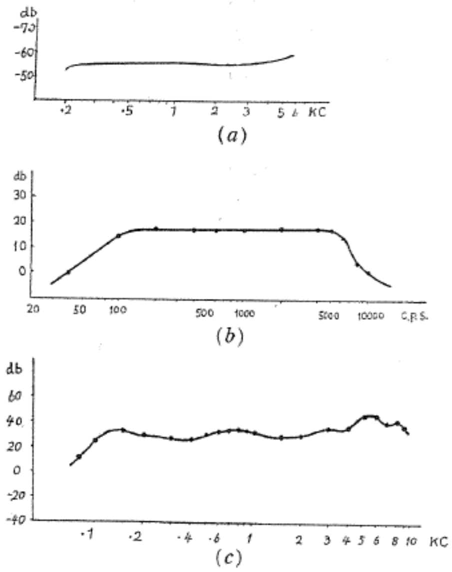
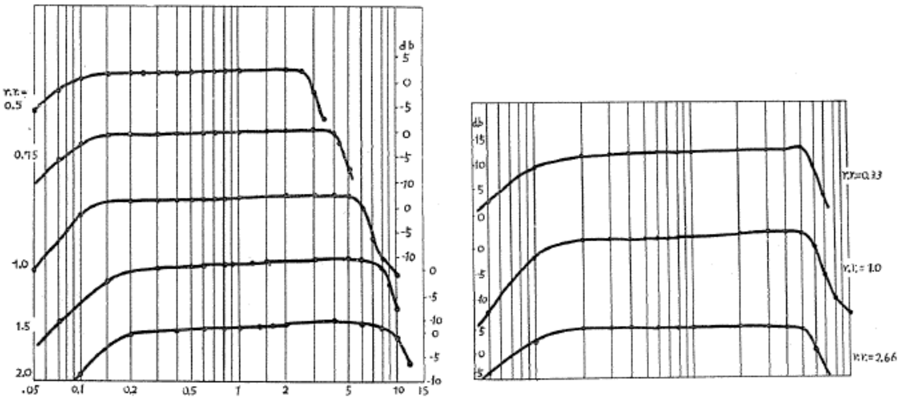


FIG. 2. Frequency response characteristics of essential parts of the systems. (*a*) Characteristic of velocity microphone. (*b*) Characteristic of tape-recorder in normal state i.e., in synchronisim. (*c*) Characteristic of loudspeaker.

Kind of distortion		D_3	D_2	D_1	N	U_1	U_2	U_3
Ratio of revolion	(n_1)	1/2	1/2	3/4	1	1 1/2	2	2
	(n_2)	1 1/2	1	1	1	1	1	3/4

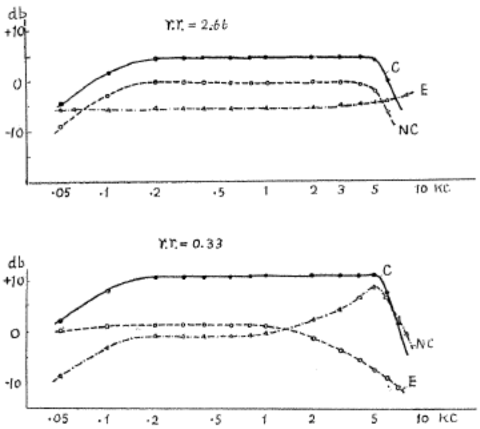
Testing condition

The change in the distortion RSD is to be defined as a ratio of the recording speed (n_1) to the reproducing speed (n_2), which we shall call hereafter ***rotational ratio*** (*r.r.*). There are two methods to change n_1 and n_2. Since the driving tork to revolve the tape-recorder is given by a synchronous motor, one can control the revolving velocity by changing the frequency of a.c. source for the motor, which we shall call frequency control method (FCM). In another way one can reach the same goal by preparing the special capstans of which sizes are different in recording from that in reproducing, which we shall call capstan changing method (CCM). The former method is preferable to the latter in the respect of smooth and continuous speed change. The latter is superior in the range of speed change, but its change is discrete and discontinuous. As we wanted to control a considerably wide extent of distortion-change in this experiment, we resorted exclusively to the latter method CCM. The former method FCM was adopted in an allowance test where the smoothness and continuity of distortion-change is vitally needed. By means of five kinds of capstan, we prepared seven steps of distortion-change shown in Table, of which the transmission characteristics are given in Fig. 3: There we see the distortions D_3 and U_3 could not be attained by making capstan size 3 : 1 from the difficulty of structural circumstances of the equipment but could be attained only by a device of a special combination of capstans as shown in the table, and, as a result, we could not but have, in these cases, considerably non-flat characteristics of which the disadvantage must be removed by insertion of equalizing networks shown respectively in Figs. 4 and 5.



(*a*) (*b*)

FIG. 3. Transmission characteristics of every testing condition of distortional step. (*a*) Characteristics of five steps ; (*b*) Two characteristics of extreme distortion-steps subjected to compensations of equalizer in comparison with normal condition.



FIGS. 4 and 5. Transmission characteristics of distortion system at the extreme deviations *r.r.* = 0.33 and *r.r.* = 2.66. *NC*: not-compensated, *E*: Equalizer used, *C*: Compensated.

Arcticulation Characteristics in General

The general articulation characteristic is given in Fig. 6. This is a mean characteristic of three observers for speeches of three talkers, given as the result

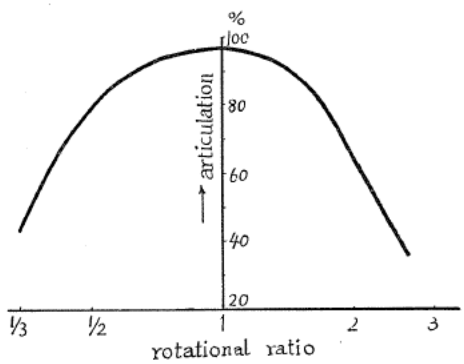


FIG. 6. General characteristic of articulation quality against the variable of rotational ratio, given by combination of voices and ears of 3 talkers and 3 listeners.

of a total of 3600 observations per condition. In this representation of the figure, the abscissa shows the variables of rotational ratio, giving as its center the condition of normal state, i.e., $r.r.=1$, and scaling both speed-up and speed-down distortions in logarithmic measure; the ordinate shows the degree of articulation quality in per cent expressed in linear scale.

Individual Characteristics per Voice

To see why this characteristic in Fig. 6 shows a little lower quality in speed-up region, i.e., in the region of $r.r.>1$, it is necessary to obtain individual characteristics per voice which we shall show in Fig. 7. By representation of this figure we can see that the two characteristics of higher pitched voices, that is, the voices of child (C) and female (F), quite coinciding with each other, come to be detached from the characteristic of lower pitched voice of male (M).

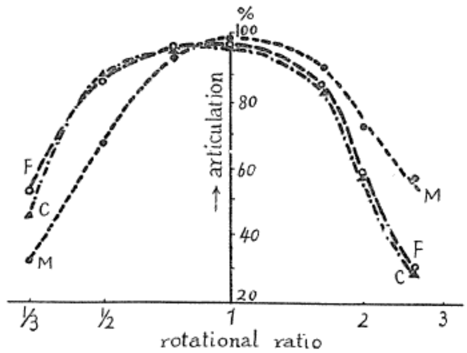


FIG. 7. Individual articulation characteristics per voice ; M : Male voice of about 130~ pitch, F: Female voice of about 200~ pitch, C: Child's voice of about 210~ pitch.

Effect of Mode of Utterance

To see the effect of mode of utterance upon the characteristics of quality response, it is sufficient to pick up only one example, i.e., the case of male voice as

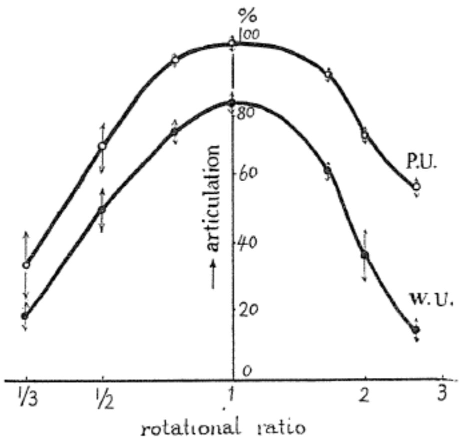


FIG. 8. Effect of mode of utterance upon the articulation characteristic in male voice, arrow-marks indicating the standard deviations of the points. PU ; Phonated utterance, WU ; whispered utterance.

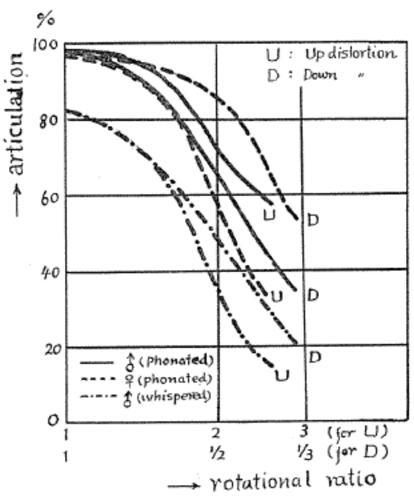


FIG. 9. Another representation of the quality responses against the distortion in two directions, showing the degree of the tendency of symmetrization of characteristics.

shown in Fig. 8. By this figure we are reminded of the fact that the symmetrization of the characteristic is more conspicuous in whispered utterance (*WU*) than in phonated utterance (*PU*). This tendency becomes more remarkable by the representation of Fig. 9, where the characteristic in the down-region is lapped over that in the up-region. More closely, in the two modes of utterance we see there the quality responses in male voice come to reverse as to the up and down distortions, but the most essential is the point that the tendency of symmetrization in *WU* is certainly remarkable.

Component Qualities of Articulation

As a sort of quality-analysis, we used to resolve syllable-unit articulation into two component-articulations, that is, vowel and consonant articulations. We show examples of such component-quality characteristics in Figs. 10 and 11. It is to be noticed: In Fig. 10 we see that not only the characteristic of consonant articulation but also that of vowel articulation bear the tendency of symmetrization in the case of whispered utterance; and in Fig. 11, i.e., in phonated utterance, only the characteristic of consonant articulation comes to be symmetrized; it is quite different with vowel articulation. These facts seem to be very important. Because they remind us of the point: Whether the characteristics of quality response can have the possibility of symmetrization as to their forms seems to depend, in large measure, on the phonal characters, that is, the phonal circumstances of speech phones towards their pitchedness. For *speech phones without pitchedness* the quality responses become symmetrized in shape. For *speech phones with pitchedness* the quality responses are short of symmetrization. In other words; articulation-quality responses nearly symmetrically in pitchless timbre, meanwhile it responses unsymmetrically in pitched timbre. As for the relation between pitchedness and unsymmetricity we cannot say anything further now.

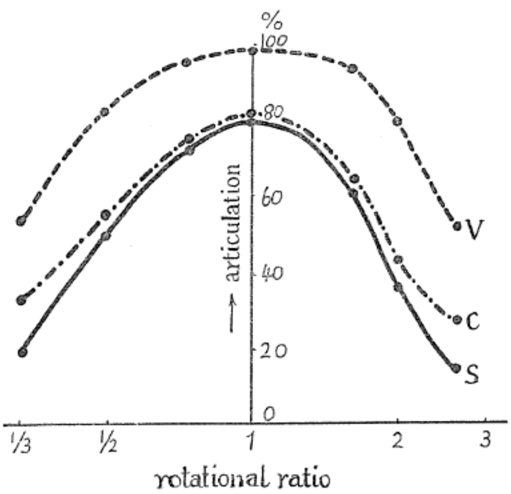


FIG. 10. Characteristics of syllable-, consonant-, and vowel-articulation in male voice in whispered utterance.

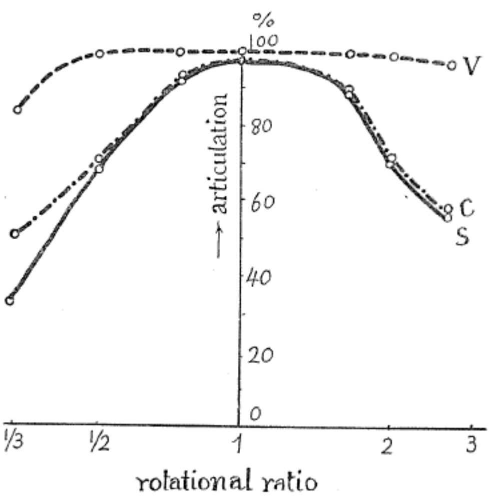


FIG. 11. Characteristics of syllable-, consonant-, and vowel-articulation in male voice in phonated utterance.

Pitch Effect on Quality Response

For accounting for the large difference between male-voice and female-voice articulation-characteristics, we must first point out the difference of pitch in two voices. We planned accordingly the following experiment: To obtain the pure change of quality-response due to the pitch change of the same voice, for example, a male talker. We show the result in Fig. 12 where the deviation of the measured quality is also given. By this experiment we can conclude roughly that there is no considerable effect caused by pitch change. By inspecting more closely we can find some effect especially in down-region where the increase of pitch by about one octave produces at most a quality-rise of about ten per cent, of which the values are yet short of accounting for the full difference of about 20 per cent due to the difference between male and female. Another explanation will be possibly needful to fill up the gap.

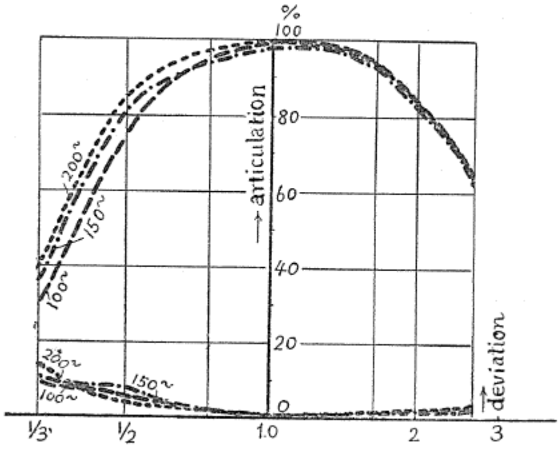


FIG. 12. Syllable-unit articulation characteristic of male voice, influenced by pitch change; showing also characteristics of standard deviations.

Articulation Study in Combined Distortion

In the previous section the transitional distortions were of nearly pure character. We treated in this section the articulation characteristics in distortion system of combined nature where rotational synchronous distortion (RSD) is accompanied by band-eliminating distortion (BED).

Experimental procedure

The recording system employed in this experiment is the same as the previous experiment (shown in Fig. 1 (*a*)). We show the reproducing system in this experiment in Fig. 13, where are inserted this time the low-pass and high-pass filters, of which the transmission characteristics are shown in Fig. 14 (*a*) and (*b*). We shall call here the former HCD (high-cut distortion), and the latter LCD (low-cut distortion). Three talkers and three listeners employed in the previous

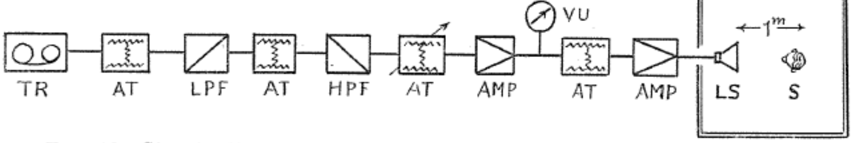


FIG. 13. Circuit diagram of reproducing system in the articulation experiment under the condition of combined distortion.

(*a*)

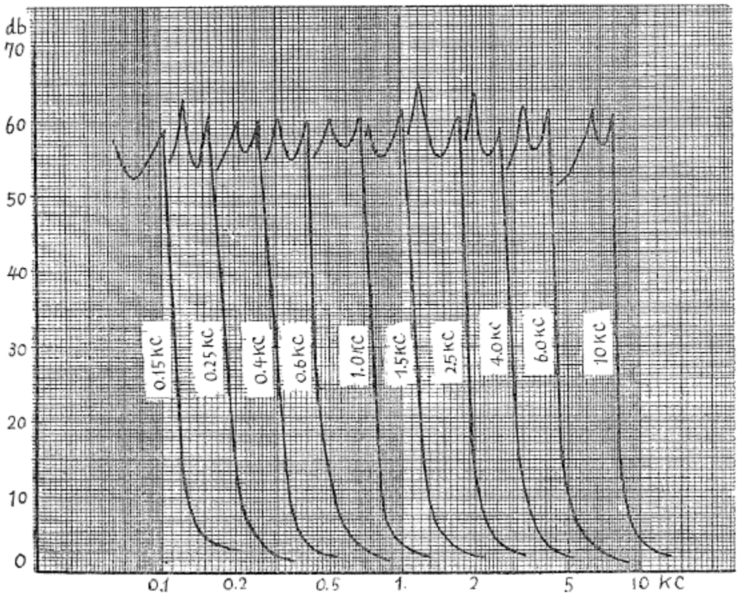


FIG. 14 (*a*). Transmission characteristics of HPF.

(*b*)

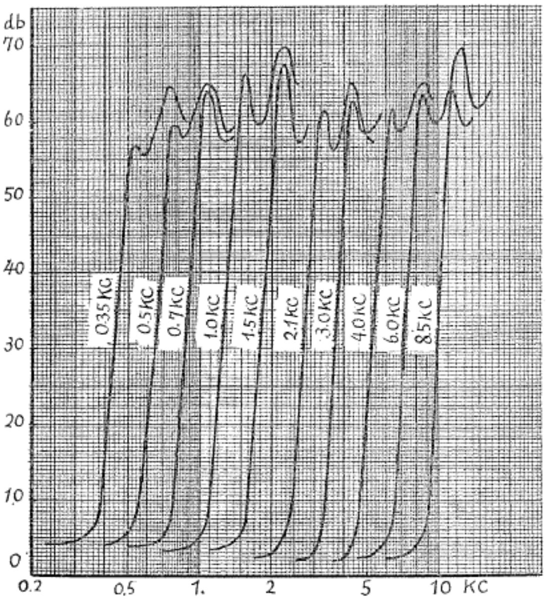


FIG. 14 (*b*). Transmission characteristics of LPF.

experiment were used here also. Logatome system is also the same as before which is very available for condition-crossing test.

Quality responses

Articulation characteristics in a combined distortion of RSD and HCD are shown in Fig. 15 (*a*) and those in the case of RSD and LCD in Fig. 15 (*b*). We can see in these figures : So long as the cutting distortions remain small the articulation characteristics are a little symmetrical (because the talking subjects are selected and combined so as to give the symmetricity to their average articulation characteristics); but the greater the cutting distortions become, the characteristics begin to move, that is, the characteristics in HCD move to the down-distortion domain (D-domain) and the characteristics in LCD to the up-distortion domain (U-domain) respectively: That is, the positions of maximum quality are departured

(*a*)

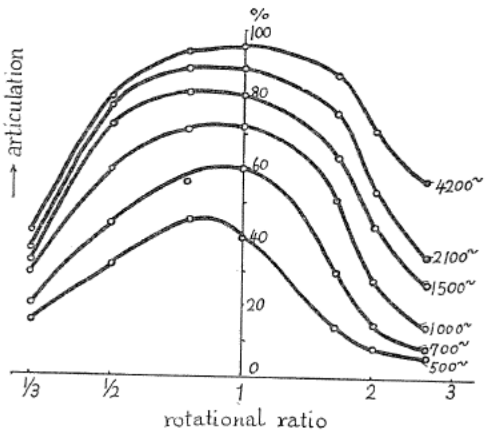


FIG. 15 (*a*). Group of articulation characteristics in a combined distortion of RSD and HCD.

(*b*)

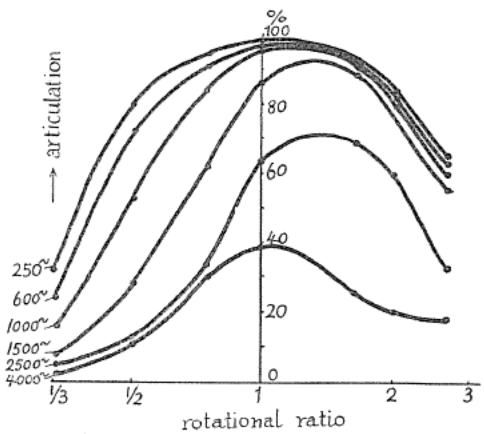


FIG. 15 (*b*). Group of articulation characteristics in a combined distortion of RSD and LCD.

from that position in normal condition. These facts mean probably that the remained quality exempt from the attack of such combined distortion seems to be a little eccentrically distributed.

By changing parameter, we can show the characteristics in Fig. 15 (*a*) and (*b*) bound together and expressed in another way as shown in Fig. 16 which is very familiar representation to us. In this figure we can trace the loci of balancing point of quality as shown by heavy chain curve. From this figure, we

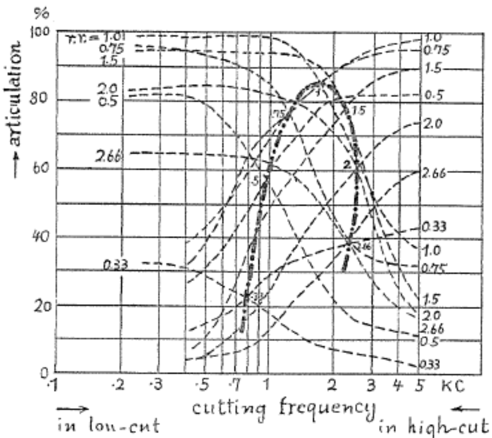


FIG. 16. High-cut and low-cut characteristics of articulation-quality responses under every seven conditions of RSD.

can lead furthermore the so-called importance (quality-distribution) curve as shown in Fig. 17. But we must point out here the following interesting fact that the importance derived from LC-characteristic is not coincide with that derived from HC-charcteristic in normal condition *r.r.* =1; Meanwhile in departured states from the normal state of synchronism, these two kinds of importance come to approach gradually, and accordingly in the extreme cases, i.e., *r.r.*=0.33 and *r.r.*=2.66, they are almost ready to attach each other. Furthermore, we must not pass over the fact that only the importances of HC-derivation move foreward in U-distortion domain, the importances of LC-derivation, on the contrary, standing still: and as for D-domain, the importances of LC-derivation move backward and those of HC-derivation preserve their inherent position given in normal state, notwithstanding the change of rotational ratio. There seems to lie so much interesting theses in quality theory. But we cannot discuss fully here. We shall study more closely on its detailed points in future.

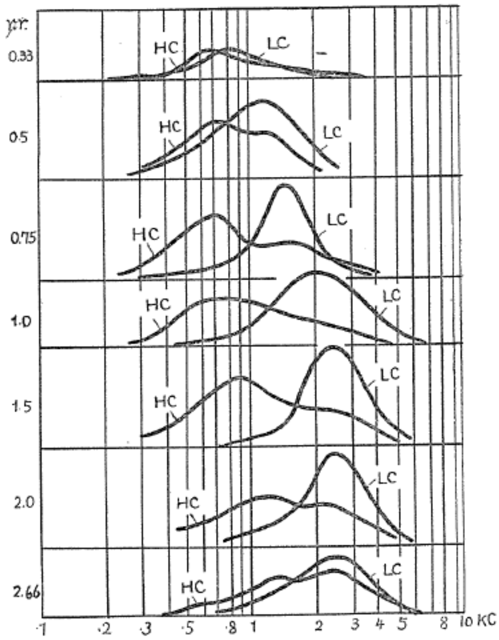


FIG. 17. Quality-distribution curves derived from Fig. 16. HC : importance derived from high-cut characteristic, LC: importance derived from low-cut characteristic.