STUDY ON DISTRIBUTION CHARACTERISTICS OF LOUDNESS QUALITY OF SUSTAINED VOWELS UNDER THE CONDITIONS OF VOCAL INTENSITY, PIANO AND FORTE

Koji MATSUURA and Junji KAWAI*

Department of Electrical Engineering

(Received May 31, 1963)

Summary: The authors pose here the problem of loudness in connection with other fundamental qualities of vowels, for example, with articulation quality which is almost independent of the influence caused by forte-piano difference in vocal intensity. One of the main characteristics of loudness is: when vowel patterns are cut down successively by Low-Cut Distortions, the difference of loudness due to vocal-intensity difference is quite conspicuous; meanwhile, when patterns are cut down by successive High-Cut Distortions, the difference of loudness caused by the same difference in vocal intensity is not noticeable. In this respect, the nature of loudness distribution in the frequency domain is exactly opposed to that of naturalness quality.

Introduction

A series of studies on loudness quality of vowels together with those on articulation and naturalness has recently been pursued as one of the fundamental stndies of speech quality[1)2)]. The problem we are here concerned with is to know how the loudness quality of vowels is distributed in the frequency domain in connection with their spectral structures; or in other words, our preesent interest is to make clear the mechanism of formation of vowel-loudness. For this purpose, it is essentially necessary to trace the relationship between the spectral deformation given purposefully and systematically and the change of the loudness distribution caused by it.

In process of the study of sevral parameters which specifies spectral pattern, the problem of uttering conditions comes in due course. For example, the spectral patterns of vowels might be subject to a specific change in shape from the difference of uttering effort. But on the other hand, this dfference has little or no effect on the articulation distribution as obviously seen in the previous researches[3)]. Thus the aim of this study naturally leads to decide whether or not the pattern variation of vowel due to uttering intensity can exert influence upon loudness distribution.

Experimentation

The experimental scheme for loudness study under the condition of vocal intensity is roughly shown by the block diagram in Figure 1. Testing materials

* Tokai Television Broadcasting Co., LTD.

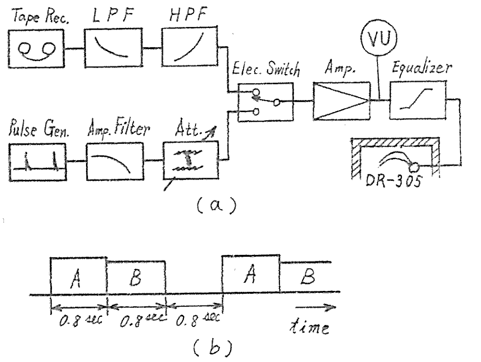

FIG. 1. (a) Block diagram of experimental scheme. (b) Signal arrangement. (A: test vowel, B: reference sound).

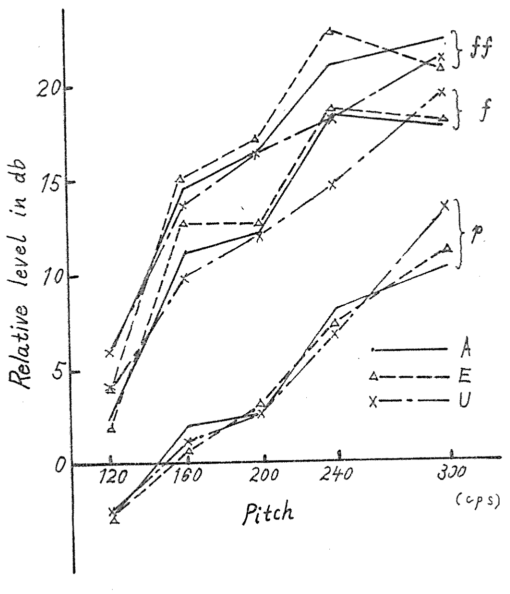

FIG. 2. "Voicing area" of the uttering subject.

were carefully collected from the voice of a male speaker* who speaks actually no provincial dialect, having some career in vocal music. The experimental conditions (such as pitch, vocal intensity, vowel kinds) were carefully examined by inspecting speaker's vocal property, *i.e.*, by measuring the "voicing area" shown in Figure 2. The pitch of 200 cps being considered most appropriate for this subject, this pitch was used throughout the experiments. As the speaker's uttering intensities, two classes of vocal effort, *forte* and *piano*, were adopted. The reasonable uttering intensities expressed in the speaker's physical output levels were also sought in the voicing area, resulting in the level difference of 10 db between *piano* and *forte*. Among five Japanese vowels three typical vowels were selected for our test; "A", a representative of wide vowels, "U", that of narrow vowels and "E", as a type of intermediate characteristics.

In this experiment, the method of measuring the loudness characteristics was employed through the successive High-Cut and Low-Cut distortions, following the same process as the precedent. As for the reference sound for the subjective loudness comparison, the non-distorted vowels used in the previous experiment gave place here to a synthesized sound of 200 cps pitch, having the spectral envelope of 6 db/octave attenuation characteristic in structure, which is almost similar to the mean pattern envelope averaged for five Japanese vowels. The method of determining the equivalent loudness under a particular condition was facilitated by using lamp signal communication between the subject and the operator, that is, the listening subject in the unechoic chamber demands by lamp indication that the operator outside should change the level of the reference sound until the subject inside judges two sounds are well matched. As the standard level of reproducing sound at the receiver, the level 5 db higher than

* F. Ishino.

the threshold of 1000 cps pure tone was appointed. The sound pressure levels of variously filtered vowels and the loudness-matched referenece sound were read, providing the measure of loudness difference between the vowels tested.

Experimental Results

We shall begin with the observations of loudness characteristic for each vowel.

Vowel "A". From the bottom representation in Figure 3, a high energy concentration in the high frequency region is evident at about 3000 cps. This specific concentration can be considered as one of the invariant formants which characterize the *vocal individuality* of the speaker. Although the speaker was instructed to utter in a sustaining manner without *vibrato*, the produced vocal sounds still represent, in some measure, the singing ones. The prominence of the invariant formant observed above seems to result from this specific uttering condition. For this vowel, under the Low-Cut distortion, the different uttering intensities give different loudness change more distinctive than for other vowels, *i.e.*, by successive Low-Cut distortion the loudness of the *piano*-utterance is easy to decrease, while that of *forte*-utterance shows a hesitation in decrease at about 3000 cps. Comparing the patterns of other vowels, it must be noticed in Figure 3 that the invariant formant at 3000 cps is cleary isolated from other formants in this vowel. As this is the case, we see here that the loundness is remarkably contributed by the existence of a particular invariant formant (2nd invariant formant).

Vowel "E". We find, in Figure 4, that the difference of uttering intensity has less influence upon the loudness characteristic. Meanwhile, it makes a much difference upon the energy distribution characteristic. Here the two peaks (energy concentrations) at 1400 cps and 3200 cps are respectively the variant formant (phonemic formant) and an invariant formant (vocal formant). This is a typical example of the pattern structure which contains several formants situated closely in the middle region of the frequency domain, and we accept that these formants are cooperative on the loudness formation.

Vowel "U". A remarkable change of pattern in *piano*-utterance apears in Figure 5; where the components in 1500-2000 cps are hardly detectable. It seems reasonable to expect that the decrease of absolute loudness level was caused by the deficiency of the higher components, and it should be noticed that this reasoning of loudness diminution is based upon the same physical correlates as those considered in the case of vowel "A" where the contribution of higher invariant formant to loudness is prominent.

After having observed the response of three vowels in loudness aspect, it can be pointed out that the fluctuating feature of balancing points of loudness characteristic curves are common to them all. The loudness balancing points in *forte*-utterance are higher than in *piano* utterance by 100-300 cps (100 cps for "A", 340 cps for "E" and 280 cps for "U". By plotting the mean characteristic curve averaged for three vowels, as shown in Figure 6, the tendency of balancing-point shift by about 180 cps from *piano*- to *forte*-condition is also observable.

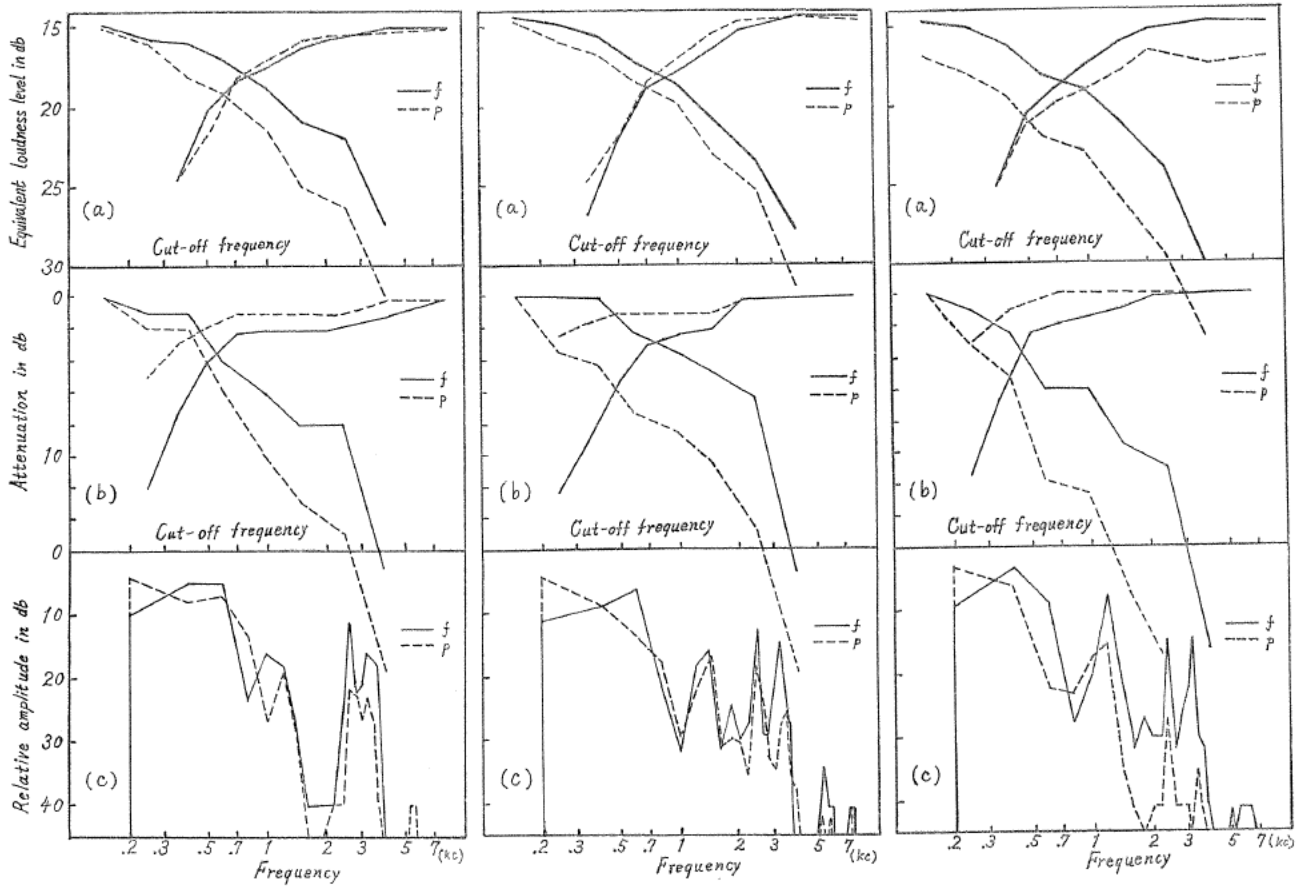

FIG. 3. Vowel "A". FIG. 4. Vcwel "E". FIG. 5. Vowel "U".

FIG. 3, 4, 5: (a) The loudness characteristic. (b) The energy distribution characteristic. (c) The spectral pattern.

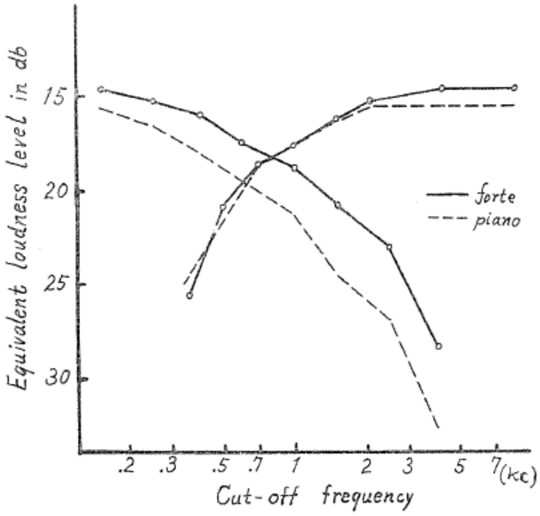

FIG. 6. Mean characteristics of loudness in *forte-* and *piano-*utterance, each averaged for three vowels.

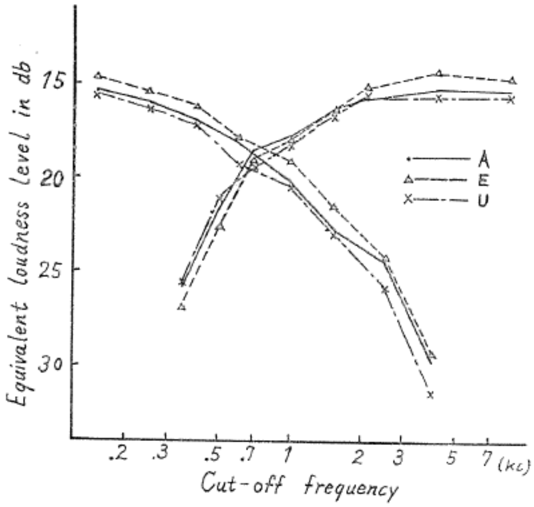

FIG. 7. Mean characteristics of loudness of three vowels, each averaged for two vocal-intensities.

From these facts, it can be said that the vocal-intensity has significant effects upon the loudness distribution.

Averaging the individual characteristic curves for vowel kind and intensity difference, we have two kinds of mean characteristic curves of loudness; one averaged for three vowels (in Figure 6), and the other for two vocal iutensities (in Figure 7), showing that the former curves have generally larger variations. A few words of explanation must be added here. The pattern properties caused by the uttering-intensity substantially differ from those caused by the vowel kind. Although the difference of patterns among vowels is characterized only by relative positions of the localized formants, these positions will be kept almost unvaried from the change of uttering intensity as easily seen in Figures 6, 7, *i.e.*, significant effect of the uttering-condition is a modification of pattern shape over all the frequency domain rather than the shift of strict positions of formants. From this and the previously observed fact that the uttering-intensity affects on the loudness more seriously than the vowel kind, it is deducible that the loudness quality is distributed dispersedly in the frequency domain. As the auditory acuities of our listening subjects contrast with each other in their over-all characteristics how these characteristics reflect upon each loudness characteristic curve is left for us to investigate as one of the interesting works in near future.

Conclusion

1. Inspecting the mean loudness characteristic for three vowels, we can conclude that the characteristic curve of the loudness quality is similar, in its form, to that of the phonemic quality rather than that of vocal quality[3]. But the quality balancing points locate lower than for other qualities, for example, at 800 cps in *forte* and at 620 cps in *piano*. (They are at about 1000 cps for phonemic quality and 1200 cps for vocal quality.)

2. Closer observation for individual characteristics tells us that in a voice having a marked upper invariant formant inherent to singing voice the invariant formant has considerable effect upon the formation of the loudness quality. In this case, due to dominant loudness factor in the upper region, quality balancing points tend, to some extent, to move towards 1000 cps. The tendency of upward-shift of balancing point is particularly conspicuous when phonemic and vocal formants are situated closely side by side as in vowel "E".

3. The center of the energy distributions (balancing points of energy distribution characteristic) fluctuates due to vocal-intensity, corresponding approximately to the location of the salient formant in each uttering condition. Meanwhile, the fluctuation of the loudness balancing points is limited in a smaller extent. Nevertheless, the effect of the vocal-intensity on balancing points should still be kept in mind; in *forte* utterance, they are higher by 200–400 cps than in *piano*.

4. In general, the loudness of vowels is affected by the vocal-intensity in a greater measure than by vowel difference. This fact is of much interest, because it means that the variation of loundness caused by the position shift of the phonemic formants according to vowel kind is retained comparatively small, meanwhile the variation of loudness due to the intensiveness of vocal formant, which is conditioned by the vocal-intensity, is much more conspicuous.

5. In addition, the difference of loudness characteristics of individual listening subjects makes an innegligible contrast with the difference in the threshold characteristics of listening subjects.

Acknowledgements

The authors are much grateful to Professor Dr. Y. Ochiai of Nagoya University for his encouragements, advice and criticism, and they are also indebted to Ass. Professor T. Fukumura for his many helps with pertinent comments and suggestion in carrying out these experiments.

References

1) Y. Ochiai et M. Oda, Sur L'intensité Sonore Subjective des Vocales Soutenues ayant les Significations Phonémique et Vocalique. MFE, Nagoya Univ., Vol. **11**, No. 1–2, Nov., 1959.
2) T. Fukumura and Y. Ochiai, Loudness Comparison between Vowels within Voice. MFE, Nagoya Univ., Vol. **7**, No. 1, May, 1955.
3) Y. Ochiai and T. Fukumura, Timbre Pattern Representation and Subjective Quality Measurement of "Forced Vocalics". MFE, Nagoya Univ., Vol. **12**, No. 1, May, 1960.