LOUDNESS COMPARISON BETWEEN VOWELS WITHIN VOICE

TERUO FUKUMURA and YOSHIYUKI OCHIAI

Department of Electrical Engineering

(Received May 18, 1955)

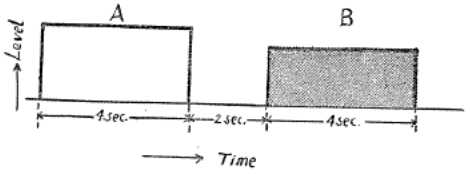
This is a part of loudness study on the so-called pitched timbre attempted on a loudness comparison between vowels within voices. Five sustained oral vowels are discriminated in respect of their loudness. It is shown that speech-sounds well matched in VU level with respect to vowel as well as voice for the purpose of timbre study of vocalic voices, are not always matched in loudness.

Introduction

It is very important to differentiate the loudness matching (LM) from the volume matching (VM) in timbre study of vocalic voices as we have already shown in a previous paper.[1] To make clear that LM and VM is not the same-thing, we must make a loudness comparison between phones well-matched in VU level. As pitched timbres for loudness matching we selected here five Japanese vowel-sounds uttered in a sustained state at a constant pitch 280~ by five calling subjects of whom two are males and two are females and one is child. There comes naturally into question matchings between voices and between vowels. As we have already obtained the vocal patterns and vowel patterns about these subjects,[2] it becomes very interesting to investigate a relationship between loudness and timbre-pattern. Such a problem of loudness viewed from timbre side is nothing more than an interpretation of timbre in terms of loudness. As a sort of *quantification* of timbre, that is, a primitive step of interpretation of timbre-quality, the study of loudness balance of speech-phones becomes more important and more instructive, apart from the practical need in the field of sound engineering.

Experiment

Since we intend to confirm the timbre study of vocalic voices in this experiment, speech-phones are selected only in sustained state. Phonemes; five oral vowels "I", "E", "A", "O", "U". Callers; male 2, female 2, child 1: each caller utters with a moderate breath-intensity, that is, *mezzo forte* at the pitch of 280~. The level of the uttered vowel-phone is about 75 db re. 0.0002 dyne/cm^2 at the position of 20 cm from the caller's lips, and the patterns of five vowels averaged for five voices are illustrated in Fig. 1. Vowel "A" of each voice (caller) is chosen as a reference vowel of loudness comparison between five vowels of the voice in question. By combining this reference vowel with each of five vowels (including "A"



A: Vowel "A" with constant level.
B: Any vowel with various level previously assigned.

FIG. 1. Illustration of signal arrangement.

itself) of the same voice to get five pairs of vowels, it enables us to try the so-called *AB* testing method. By picking up the pure stationary part out of a caller's sustained vowel "A" of about four seconds duration and recording it on tape recorder, the signal *A* is thus obtained. With the pause-interval of about two seconds, any one vowel of the same caller, treated in the same way as before, is recorded as the signal *B*: these are illustrated schematically in Fig. 2. The recording and reproducing levels of the signal *A* are kept always constant through all the voices, while those of the signal *B* are stepped by 3 db in the level-range of 12 db around that of the signal *A*. Arranging the level height of the signal *B* in random order in time, a succession of pairs of the signals *A* and *B* which are to be compared with each other, is obtained, and is led to the listening subjects by such a scheme as shown in Fig. 3. The level of the reproduced signal *A* is about 65 db re. 0.0002 dyne/cm^2 at the position of the listening subjects. The fidelity of the electrical and acoustical equipments used is most reliable over the frequency range from 0.1 to 6 kc. The listening subjects, 4 young students, were previously tested about their hearing acuity by a measurement of auditory threshold and they were considered quite acustomed to such a discriminatory judgment as loudness comparison after an excercise of loudness balance between pure tones. They are now instructed to vote when they judge that the signal *B* is louder than *A*; judgment is repeated ten times for each combination of the signals *A* and *B*. Though the judgment criteria of "equality" and "less loudness" may also be established for such a loudness comparison, we used here exclusively "louder" criterion for simplicity.

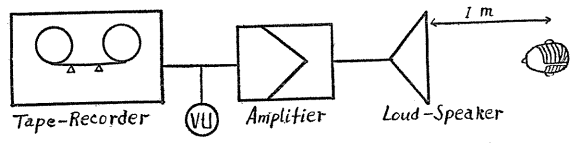


FIG. 2. Block diagram of experimental scheme.

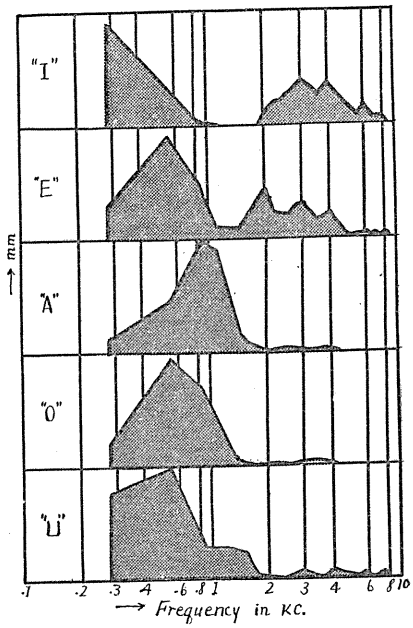
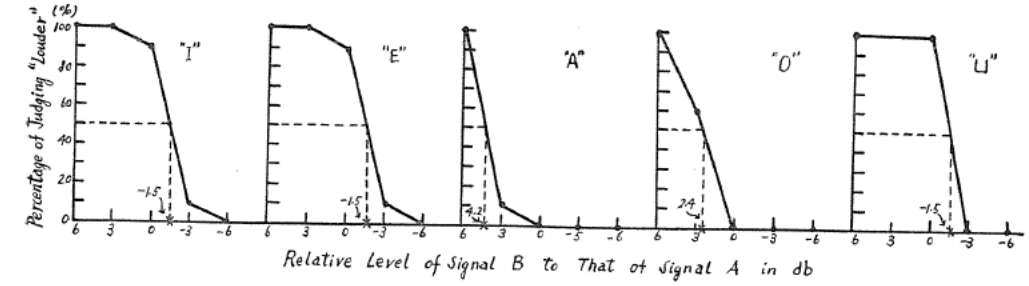


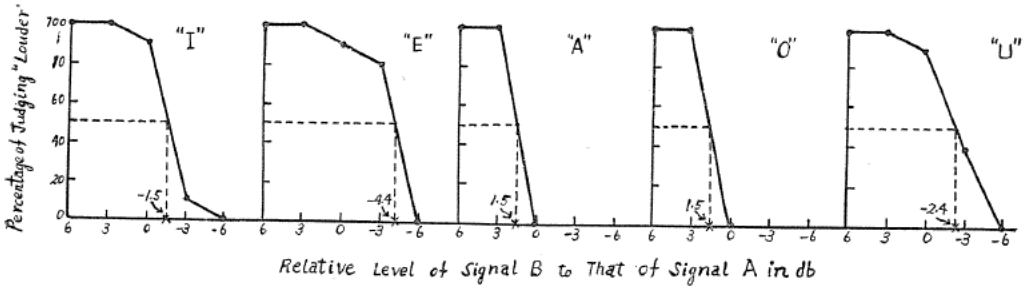
FIG. 3. Formant-patterns of five vowels averaged for five voices (callers). Pitch; 280~, uttering breath-intensity; *mezzo forte*.

Experimental Results

The results of this testing method for two combinations of calling and listening subjects are shown in Fig. 4, where the abscissa stands for the level of the signal *B* with reference to that of the signal *A* and the ordinate gives the percentages of the subject's louder judgment for the signal *B* of a corresponding level. From these curves, we intend to obtain a level difference that just produces the loudness difference between the signal *A* (vowel "A") and the signals *B* (any one of five vowels, of course, including "A"), that is, the value on the abscissa corresponding



Voice: S. M, (Male), Listening subject: Nag. (Male).



Voice: Y. H. (Female), Listening sudject: Hat. (Male).

FIG. 4. Graphycal representation of the results of loudness comparison between vowel "A" and each of five vowels (including "A") by means of *AB* testing method. These are the examples for male and female voices. The listening subjects, Nag. and Hat., are arbitrarity chosen from four subjects as their representatives.

to the ordinate of 50% is taken as a threshold of louder judgment (*louder threshold* in abbreviation) of any vowel to the reference vowel "A". Thus, threshold values of five vowels of each voice are determined for each listening subject: in all, one hundred values of such a threshold are obtained. As they do not show so much large discrepancy between listening subjects, they are averaged for them to reduce the concerning factors only to two, namely, voice and vowel. To the values thus averaged, a statistical treatment of variance-analysis is then applicable, and the results are given in Table. Comparing with error term, vowel term shows a more

TABLE. Results of Variance Analysis

Variance due to	Degree of freedom	Sum of squares	Unbiased estimate
Voice	4	11.0	2.8
Vowel	4	79.6	19.9
Error	16	29.8	1.9

significant large variance. This means that the level differences needed to produce a louder impression than the reference vowel are not the same among the five vowels; in other word, the discriminatory thresholds of loudness between vowel-phones can be determined within a voice and their values differ considerably according to the kinds of vowel; then it remind us of the fact that there may be some relations between such threshold value and timbre-pattern. As we have seen in the Table, the voice term shows small variance: this enables us to say that the discriminatory thresholds between vowels above mentioned have the fairly similar

tendency in spite of the difference of voices. Here we return again to Fig. 4, where the signals B of abscissa of 0 db are matched in volume level with the signal A, and it is seen that their corresponding percentages of louder judgment show a considerably large variety with respect to the kind of vowel. Averaging for all voices and listening subjects about this value of louder judgment for the volume-matched signals, we obtain the results as shown in Fig. 5, where the kind of vowel is shown on the abscissa and the ordinate represents the percentages of louder judgment. From this figure we can see that the vowels "A" and "O" are seldom to be judged as louder when they are volume-matched with the reference "A", and the voting frequencies of listening subjects, can be considered in this case as statistical error; while the vowels "I", "E", "U" are very frequently judged as louder, even when they are volume-matched with "A". This means that there is definite loudness difference between vowels matched in volume level, and also tells us that the vowel-phones, under the condition VM, can be discriminated by the loudness quality, reflecting somewhat characters of their timbre structure. Averaging for all voices about the threshold values, the curve to represent the loudness relationship between vowels is obtained as shown in Fig. 6. There we can see that the vowel "A" is to be gained by 2.6 db in volume level to be judged as just louder than itself. On the contrary, the vowel "I" must be attenuated by 1.5 db to become just louder than "A". . . . The distinct result obtained from this figure is that the vowel-phones are classified into two groups in such loudness-threshold aspect; namely, "A", "O" belong to a softer group and "I", "E", "U" to a louder group. This fact is also remarkable in Fig. 5, meaning that the vowel-phones have only two classes of quality in their loudness aspect, though they have five qualities in their phoneme aspect; in other words, the timbre-

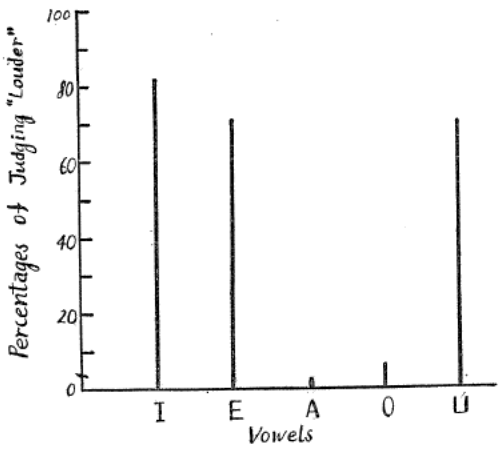


FIG. 5. Percentages of louder-judgment well matched in VU with the reference vowel "A".

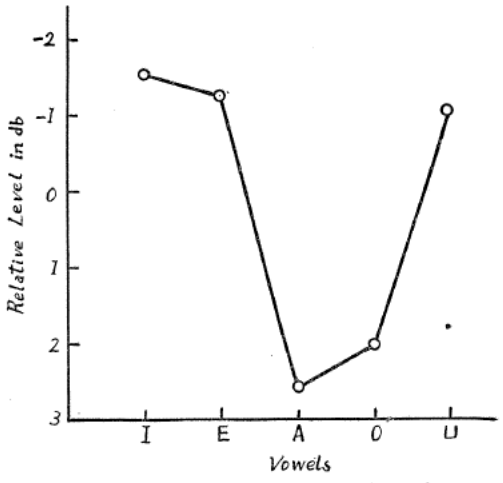


FIG. 6. Curve representing the relationship between vowels. On the ordinate, the positive sign means the amplification and the negative one means the attenuation, needed to just discriminate each vowel from vowel "A" in respect of loudness.

patterns of vowel can be interpreted by loudness quality as having lower classifications than those by phoneme quality.

Summary

We have represented the loudness relation of vowels well-matched in volume level by the values of louder-threshold in a judgment of loudness comparison. By such a method we can classify all five Japanese vowels in two divisions with respect to their loudness. One, a group of vowels "A" and "O": the other, a group of vowels of "I", "E" and "U". It is to be noted that the loudness of "A", "O" group is less than that of a group of "I", "E", "U", provided that every vowel in question is quite matched as to volume level. From a variance analysis we can conclude that there are many problems in loudness between vowels (or between phonemes). The problem in loudness between voices is left for future investigations.

References

1) Y. Ochiai, S. Saito, M. Watanabe and T. Fukumura; Quality Studies of Speech Transmission System especially with Reference to Characteristic Forms in Frequency Response. Memoirs of the Faculty of Engineering (M.F.E.), Nagoya Univ., Vol. 6, No. 2, Nov. 1954.

2) Y. Ochiai and T. Fukumura; Timbre Study of Vocalic Voices. M.F.E., Nagoya Univ., Vol. 5, No. 2, Oct. 1953.

3) T. Fukumura and K. Hara; Phoneme Figures of Sustained Oral Vowels by Two Dimensional Representation. (I) and (II). M.F.E., Nagoya Univ. Vol. 6, No. 2, Nov. 1954.