ISSN 0287-5357

SOPHIA LINGUISTICA

Working Papers in Linguistics

Number 13
1983

The Graduate School of Languages and Linguistics
Linguistic Institute for International Communication
Sophia University
Tokyo 102, Japan

Academic Press, 1974), p. 302.

40 Morris Halle, "Theoretical Issues in Phonology in the 1970's," in *Proceedings of the 7th International Congress of Phonetic Sciences* (Montreal, 1971), cited in Sommerstein, *op. cit.*, p. 106.

41 *loc. cit.*

42 *op. cit.*, p. 97.

SIMILARITY DATA IN THE MEASUREMENT OF OPTIMAL FREQUENCY BANDS FOR VOWELS

Kikuo Maekawa

Summary

A new procedure to determine the Optimal Frequency Band of vowels is proposed on the basis of the criticism of the traditional method. In the new method, optimal frequency band is determined as a simple function of the psychological similarity we perceive between the filtered stimuli and the original non-filtered stimulus. Some favorable characteristics of the new method are discussed in connection with the result of its application to the analysis of the two Japanese vowels/i/and/a/.

Introduction

The measurement of the optimal frequency band (henceforth OFB) of a given speech sound —notably the vowel— is of great importance in the research of the Verbo-tonal System both in theory and in practice. However, the traditional method of measuring OFB has considerable defects when looked at from the theoretical point of view. The aim of this paper is twofold: to propose a new method of measuring OFB and then to show the result of its application in the analysis of some Japanese vowels.[1]

Problems

The OFB of a given speech sound can be defined, in informal terms, as the frequency band through which the filtered sound becomes most similar to the original non-filtered sound in phonetic quality.[2] Experimental workers of the Verbo-tonal System have made it a custom to determine the OFB as the function of the identification rate in perception tests. A set of filtered speech sounds is presented, in random order, to the subjects (who are most commomly the native speakers of the language under consideration) and they are requested to identify the phonological attribute of each filtered stimulus according to their own intuition. The OFB of the original sound is determined usually as the band of the filter used for the stimulus which obtained the highest identification rate. At least two defects can be pointed out concerning such experimentation, however. Firstly, we cannot determine, even in principle, the OFB of the foreign speech sounds in this way, for it is naturally expected that the subjects who are able to identify —correctly or incorrectly— the phonological attribute of the foreign speech sound have conciderable amount of experience in hearing that foreign language. What really we want to know is the reaction of naive listeners who, desirably, hear that foreign sound for the first time. In this connection, Landercy and Renard called the OFB of the mother tongue and of the foreign languages, respectively, *l'optimale absolue* and *l'optimale de la faute*.[3] If necessary, we will distinguish the two as OFBA (Absolute) and OFBF (Fault).

The second defect is a more practical one and holds in the case of OFBA as well. As a result of the identification experiment, it happens sometimes that more than two filtered stimuli obtain the same identification rate (often 100%), though there exist accessible differences in their phonetic qualities. A good example can be cited from Takata's experiment with five Japanese vowels.[4] In his experiment, using the Japanese /a/ vowel, three filtered stimuli obtained 100% identification rate (among eight stimuli), and the frequency bands used for these stimuli were (a) 600-1200 Hz, (b) 800-1600 Hz and (c) 1200-2400 Hz. The Japanese male /a/ vowel sounds like IPA [a] when filtered through band (a), and like IPA [æ] through (c). Whatever the exact phonetic qualities of these stimuli, what matters here is that the differences among these stimuli —which should have been perceived by most of the subjects— were not reflected in the identification rate at all. Similar examples can be found in the experiments of Landercy and Renard with French oral vowels.[5]

As this kind of defect is clearly correlated with the psychological nature of the identification process itself, we cannot but adopt a way of data collection other than identification to overcome, even partially, the problems mentioned above. The following experiments are designed to examine the empirical validity and efficacity of the use of similarity data in measuring the OFBA of two Japanese vowels, viz, /i/ and /a/.

Experiment

Two Japanese male speakers pronounced the five Japanese vowels several times in an anechoic room associated with the Phonetics Laboratory of Sophia University. Two tokens of /i/ and /a/ which seemed to be the most natural were chosen from the recorded materials.[6] These tokens are filtered with Bruel & Kjoer Spectrumshaper Type 5612 connected to a Measuring Amplifier Type 2607. The frequency bands used in the experiment are shown in Table 1 below; these frequency bands are almost identical to the bands used in Takata's experiment. Henceforth we will distinguish the stimuli by means of the subscripts as shown in the leftmost column of the table. If necessary, we will refer to the original sounds as i_0 and a_0.

Table 1 Bands of frequency used

Stimulus	Band of frequency
i_1	630 – 1250 Hz
i_2	800 – 1600 Hz
i_3	1250 – 2500 Hz
i_4	1600 – 3200 Hz
i_5	2500 – 5000 Hz
i_6	3200 – 6400 Hz
a_1	316 – 630 Hz
a_2	400 – 800 Hz
a_3	630 – 1250 Hz
a_4	800 – 1600 Hz
a_5	1250 – 2500 Hz
a_6	2500 – 5000 Hz

Frequencies in the table denote the center frequencies of the 1/3 octave band-pass filters contained in Bruel & Kjær Spectrumshaper.

These fourteen stimuli (including i_0 and a_0) are submitted to a traditional identification test on the one hand, and the psychological similarities between the original sounds and the filtered stimuli are measured by means of the paired comparison method on the other. In the latter test, all ${}_6P_2$ (= 30) pairs of the filtered stimuli are presented, in random order, to the subjects and they are asked to determine which stimulus of the pair is more similar to the original sound recorded immediately before the presentation of each pair. The schematic representation of the paired comparison procedure is shown in Fig. 1.

FIG. 1 THE PROCEDURE OF THE PAIRED COMPARISON TEST

N // W. N. // X_0 X_i X_j // ·········

N : Number of the pair (from' first' to' thirtieth' pronounced by male voice).
W. N. : White noise five seconds long.
X_0 : The original non-filtered sound (i_0 and a_0).
X_i ,X_j : The filtered stimuli ($i, j = 1, 2, \cdots, 6$; $i \neq j$).
······ : six second interval preceding next pair.

Sixteen subjects —eight males and eight females— participated in the identification tests and in the paired comparison for /a/. Only forteen of them took part in the paired comparison test for /i/.

Results

The results of the identification tests are shown in Tables 2 and 3 below. In the last row of each table are shown the rates of correct identification for the six

Table 2. Result of identification test for Japanese vowel /i/.

	i_1	i_2	i_3	i_4	i_5	i_6
Subject 1	a	a	u	i	i	i
Subject 2	a	a	i	i	i	i
Subject 3	a	a	i	i	i	i
Subject 4	a	a	i	i	i	i
Subject 5	a	a	i	i	i	i
Subject 6	a	a	i	i	i	i
Subject 7	o	a	i	i	i	i
Subject 8	o	a	i	i	i	i
Subject 9	u	u	i	i	i	i
Subject 10		a	u	i	i	i
Subject 11	o	a	i	i	i	i
Subject 12	a	a			i	i
Subject 13	o	e	i	i	i	i
Subject 14	u	i	i	i	i	i
Subject 15	a	a	i	i	i	i
Subject 16	a	a		i	i	i
%	0	7.1	71.4	92.8	100	100

In the last row of the table, we show the rates of correct indentification.

Table 3. Result of identification test for Japanese vowel /a/.

	a_1	a_2	a_3	a_4	a_5	a_6
Subject 1	o	o	a	a	u	i
Subject 2	o	o	a	a	a	i
Subject 3	o	o	a	a	a	i
Subject 4	o	o	a	a	a	i
Subject 5	a	o	a	a	e	i
Subject 6	o	o	a	a	u	i
Subject 7	o	o	a	a	a	i
Subject 8	o	o	a	a	u	i
Subject 9	o	o	a	a	a	i
Subject 10	o	o	a	a	a	i
Subject 11	o	o	a	a	a	i
Subject 12	o	o	a	a	a	i
Subject 13	o	o	a	a	a	i
Subject 14	o	o	a	a	u	i
Subject 15	o	o	a	a	a	i
Subject 16	o	o	a	a	e	i
%	6.25	0	100	100	62.5	0

In the last row of the table, we show the rates of correct identification.

filtered stimuli. Some of the elements of the tables remain unspecified because the subjects were not forced to give the answer when they felt considerable difficulty in identifying the stimulus in terms of Japanese vowels or in terms of Japanese syllables.

The results of the paired comparison are summed up in the following manner: when we submit n stimuli $(s_1, s_2, \cdots\cdots, s_n)$ in paired comparison, and when N subjects participate in the test, we have ${}_nC_2 \cdot N$ judgments at the end of the test. We will represent the judgment of the i-th subject of the pair s_j and s_k as

$$e_{ijk} = \begin{cases} 0: \text{when subject's judgment is } s_k > s_j, \\ 1: \text{when subject's judgment is } s_j > s_k, \end{cases}$$

$$(1 \leqq j \leqq n \text{ and } 1 \leqq k \leqq n \text{ and } j \neq k).$$

Here, $s_j > s_k$ means that s_j is more similar to the orginal sound than s_k. In addition, it will be natural to assume the following relation, viz,

(1) $e_{ijk} + e_{ikj} = 1.$

we can thus obtain the frequency of the judgment $s_j > s_k$ according to the following formula,

$$f_{j>k} = \sum_{i=1}^{N} e_{ijk} \qquad (j \neq k)$$

And by the relation (1) assumed above, we see that

$$f_{j>k} + f_{k>j} = \sum_{i=1}^{N} (e_{ijk} + e_{ikj}) = N$$

Once the frequencies of the judgments have thus been decided, it will be convenient to show them up in the form of a matrix as in Tables 4 and 5. In these matrices, the elements of the j-th row in column k represents the frequency $f_{j>k}$ defined above. The diagonal elements in the matix, in other words the frequencies of the judgments when s_i and s_j are the same, are calculated theoritically. By assuming the *a priori* probability of 0.5 for every judgment and by the relation (1) abore, we have

$e_{ijk} = e_{ikj} = 0.5$ (where $j = k$),

from which we have

$$f_{j>j} = \sum_{i=1}^{N} e_{ijk} = N/2$$

where $f_{j>j}$ is the diagonal element. The elements in the last row of Tables 4 and 5 —noted as $F_{j>\cdot}$ $(1 \leqq j \leqq n)$—indicate the sum of each columm, viz,

$$F_{j>\cdot} = \sum_{k=1}^{n} f_{j>k}$$

The $F_{j>\cdot}$ thus represents the total frequency of the judgments in which the stimulus s_j is judged to be more similar to the original sound than the other stimuli. Finally, the following remark is to be added. Every pair of two given stimuli having been presented twice in our experiments, there were some cases

Table 4. Result of paired comparison for japanese vowel /i/.

	i_1	i_2	i_3	i_4	i_5	i_6
i_1	7	10	13	13.5	14	14
i_2	4	7	13.5	13	14	14
i_3	1	0.5	7	12	12.5	10.5
i_4	0.5	1	2	7	10	7
i_5	0	0	1.5	4	7	3.5
i_6	0	0	3.5	7	10.5	7
$F_{j>}.$	12.5	18.5	50.5	56.5	68	56

Only fourteen subjects participated in this test.

Table 5. Result of paired comparison for japanese vowel /a/.

	a_1	a_2	a_3	a_4	a_5	a_6
a_1	8	9.5	16	16	14	1.5
a_2	6.5	8	16	15	15	1
a_3	0	0	8	8.5	1.5	0
a_4	0	1	7.5	8	1	0
a_5	2	1	14.5	15	8	0
a_6	14.5	15	16	16	16	8
$F_{j>}.$	31	34.5	78	78.5	55.5	10.5

where some subjects gave contradictory answers. For example, a subject might give the judgment $s_j > s_k$ in one of the two judgments, and in the other presentation of the same set of stimuli he might give the judgment $s_k > s_j$. In such cases, I made it a custom to assign the number 0.5 to both e_{ijk} and e_{ikj}.

Interpretation

This section will be concerned with the statistical examination of the results shown in Tables 2, 3, 4, and 5. A fundamental problem here is to know whether we can bypass, by the use of the similarity data, two defeciencies of the identification data discussed in section 2.

Because the two sets of data differ with respect to their mathematical nature of scaling —the identification data is on an interval scale whereas the similarity data is on an ordinal scale— I would like to begin by treating all of them as data on an ordinal scale.[7]

In the case of the /i/ vowel, the relative order of the six filtered stimuli determined as the function of the identification rates is:

$i_5 = i_6 > i_4 > i_3 > i_2 > i_1$

while the order determined as the function of the $F_{j>}.$ in Table 4 is:

$$i_5 > i_4 > i_6 > i_3 > i_2 > i_1$$

The rank order correlation coefficient (Spearman's ρ) calculated from the two orders being equal to 0.899, we can reject the null hypothesis of $\rho = 0$ in the population at the 0.01 level of significance. On the other hand, the relative order of the six filtered stimuli of the /a/ vowel is determined as

$$a_4 = a_3 > a_5 > a_1 > a_2 = a_6$$

from the identification rates, and as

$$a_4 > a_3 > a_5 > a_2 > a_1 > a_6$$

from the $F_{j>i}$ in Table 5. In this case, the rank order correlation coefficient is equal to 0.883 and the null hypothesis can be rejected at the 0.025 level of significance.[8] Statistically speaking, the two independent experiments are closely correlated—though not totally identical—with respect to the relative order of the stimuli determined by the experiments.

Moreover, two important facts should be pointed out from the theoretical point of view of OFB. First, the stimuli judged to be most similar to the original sounds (i_5 and a_4) always obtained the highest rates (100%) in the identification tests. Secondly, the stimuli of the highest similarity are always determined uniquely: there is no "ties" in the order determined by means of the similarity, but not with respect to the identification data. To put it differently, we can say that the differences of phonetic quality existing between i_5 and i_6, on the one hand, and between a_3 and a_4 on the other hand, are correctly reflected in the similarity data. It should be noted also that the stimuli of the highest similarity are exactly the same as the stimuli chosen in Takata's experiment, though he drew his final coclusion not only by means of the identification rate but also by reffering to the sonority of the filtered sounds.

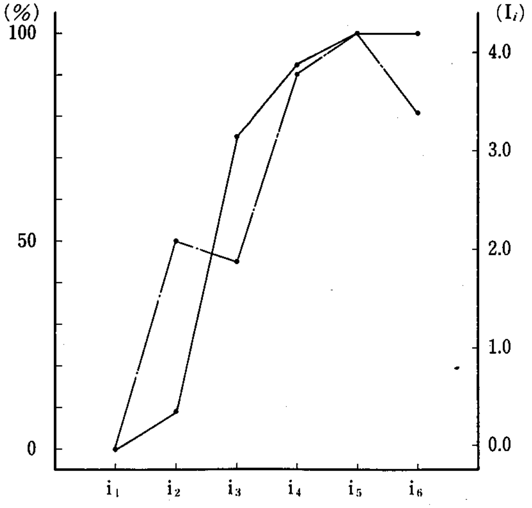

Fig. 2. Interval scale for /i/ (I_i) and rates of identification.

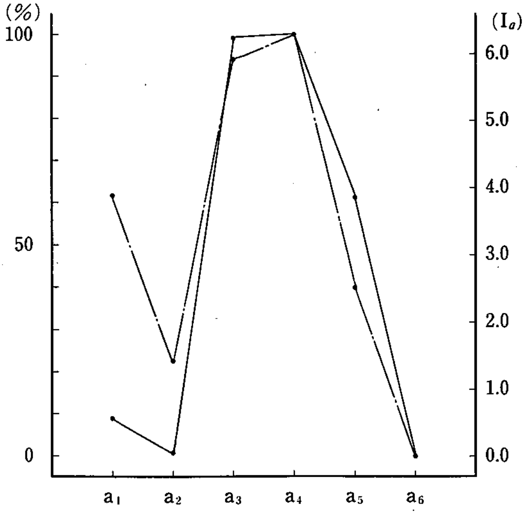

Fig. 3. Interval scale for /a/ (I_a) and rates of identification.

In Fig.s 2 and 3, the real line represents the rates of identification and the broken line represents the points of each stimuli on the newly constructed interval scales from similaity data I_i and I_a.

An alternative way of comparing our two results consists in the making up of an interval scale from the similarity data. To achive this objective, I have applied L.L. Thurstone's *law of comparative judgment* (caseV) to our similarity data.[9] The two interval scales thus constructed are graphically compared with the identification interval scales in Fig.s 2 and 3. The two important characteristics of the simiarity data just mentioned remain invariant after the application of the *law of comparative judgment.*

Conclusion and Further Problem

From the foregoing experiments, we see that the similarity data are fairly useful in order to decide the OFBAs of given vowel sounds of the mother tongue, especially when there are more than two filtered stimuli resembling closely each other.

However, one important problem still remains to be resolved in this respect: the quantitaive evaluation of the differences of the phonetic quality captured in the similarity data. To the best of my knowledge, this problem could be tested non-parametrically only by increasing the number of the presentation of the same sets of stimuli.

This measurement technique will be of great value if we try to measure directly the OFBF of foreign vowel sounds, which is still only a theoretical construct in the Verbo-tonal System.[10]

Notes

1 This paper is a revised English version of my 1981 manuscript "Efficacité empirique des données de similitude dans le cas de la mesure des bandes optimales absolues." All experimental works reported here were carried out mostly in the summer of 1981 as a pilot study for my 1982 MA thesis (See, note. 10 below).

2 The definition of OFB differs somewhat depending on the author. The definition adopted here is based upon the one shown in the following article: Petar Guberina, "Les appareils Suvag et Suvaglingua," *Revue de phonêtique appliquée,* 27-28 (1973), 14.

3 Albert Landercy and Raymond Renard, "Perception des voyelles françaises filtrées," *Revue de phonêtique appliquée,* 32 (1974), 11-32.

4 Yōji Takata, "Tonalities of Japanese speech sounds for the establishment of some Japanese word lists for audiometric purposes -- based on Verbo-tonal audiometry," MA thesis, The Graduate School of Languages and Linguistics, Sophia University, Tokyo, 1979.

5 Albert Landercy and Raymond Renard, "Perception," (1974). ———, "Zones frêquentielles et reconnaissance de voyelles françaises," *Revue de phonêtique appliquée*, 33-34 (1975), 51-79.

6 These two vowels were pronouced by two different speakers as the result. In addition, the vowels —/i/ and /a/— were chosen as the representatives of diffuse vowels and compact vowels respectively.

7 For the discussions concerning the levels of scaling, see, for example, Yoshihisa Tanaka 田中良久, *Shinrigakuteki sokuteihō* 心理学的測定法 (Psychometrics), (Tokyo: Tokyo University Press, 1977).

8 The results of the statistical test on Spearman's ρ discussed here are not exact, because the orders decided by the rates of identification contain some "ties" which is not assumed in the mathematical foundation of the non-parametric test for the Spearman's ρ. The null hypothesis, however, can be rejected at least at the 0.05 level of significance in all possibilities.

9 L.L. Thurstone, "A law of comparative judgment," *Psychological Review*, 34 (1957), 273-286. See, also Y. Tanaka (1977) above.

10 A tentative way to measure the OFBF by means of the technique reported here is shown in my 1982 MA thesis, "Mesure psycho-acoustique des bandes optimales des voyelles françaises /y/ et /ϕ/ telles que perçues par les japonophones," The Graduate School of Languages and Linguistics, Sophia University, Tokyo, 1982.