Bulletin of the Faculty of Engineering,
Hokkaido University, No. 77 (1975)

On the Phonetical Classification of Nasals and Liquids

Masaru SHIMBO*

(Received March 31, 1975)

Abstract

The present paper is an extension of the theoretical classification of vowels based on the information-theoretical point of view to nasals and liquids, both of which have more formant characteristics than other consonants. Calculating parameters from the formant frequency data of Swedish nasals and liquids, and plotting them on the phonetical vowel plane, we can classify these phonemes by the indices in the information-theoretical structure of sensation. Though the present data are limited and few, we can see that nasals as well as liquids are divided into several groups by the ensemble of the foregoing indices, located among normal and modified vowels.

1. Introduction

Here we summarize the analysis of experimental data of Swedish nasals and liquids based on the theoretical classification of vowels.

The theoretical classification of vowels by the indices of the information-theoretical structure of sensation was first suggested by K. Kondo[1], and similar expositions followed along the same lines of thought: Ordinary normal[2], modified[3),4] and neutral vowels[4] are classified by the principle. The present paper is an extension thereof to nasals and liquids, both of which have more formant characteristics than other consonants. Since an application to liquids has only been predicted in the previous papers[3], the empirical and/or experimental confirmation of liquids as well as nasals are mainly dealt with in this paper. The methods of analysis are the same as those used in the foregoing analysis. The formant frequencies of Swedish nasals and liquids have been aquired from G. Fant's analysis[5].

2. Structure of Sensation

The theoretical classification of principal vowels by the indices from the standpoint of the information-theoretical structure of vowel sensation is discussed in the previous papers[1)~4]. It is based on the fundamental recognition that the increment of the information-theoretical entropy function of a stimulus system gives a measure of sensation. It will briefly be recapitulated as follows.

A discrete set of information elements is considered and P out of the total number M of these are selected. Such a system has a monostimulus character with a single variable P, while M is thought to be fixed. In order to represent

* Division of Information Engineering, Graduate School of Engineering, Hokkaido University, Sapporo 060, Japan.

the information of the system, we introduce the concept of entropy as a measure from the information-theoretical point of view. It is defined by

$$H=\log\binom{M}{P}, \tag{2. 1}$$

where

$$\binom{M}{P}=\frac{M!}{P!\,Q!},$$
$$M=P+Q. \tag{2. 2}$$

Formula (2. 2) is called the first compensation relation. Approximately, (2. 1) is given by

$$H=M\log M-P\log P-Q\log Q \tag{2. 3}$$

by Stirling's formula for sufficiently large numbers of P, Q and M.

It should be noted that no information is communicated when no change takes place. In other words, the communicated part of an ensemble of informations is contained entirely in the increment of entropy, which means

Definition 1. The increment of an entropy function is the measure of sensation.

The sensation of the system, the entropy of which is defined by (2. 3), is then given by

$$S=\varDelta H=\varDelta M\log M-\varDelta P\log P-\varDelta Q\log Q \tag{2. 4}$$

with a supplementary condition

$$\varDelta M=\varDelta P+\varDelta Q, \tag{2. 5}$$

where $\varDelta H$, $\varDelta P$, $\varDelta Q$ and $\varDelta M$ are the increments of H, P, Q and M. Formula (2. 5) is called the second compensation relation. Hence the lemma:

Lemma 1. The sensation of a monostimulus system is represented by the linear combination of three logarithmic terms with the second compensation relation.

The equation (2. 4) agrees in structure with the Weber-Fechner criterion. By substitution of (2. 5) in (2. 4) it can also be written in the form

$$S=\kappa(\log p^m-\log q^n), \tag{2. 6}$$

where

$$p=\frac{P}{M}, \qquad q=\frac{Q}{M},$$
$$\kappa=-\frac{\varDelta P}{m}=\frac{\varDelta Q}{n}. \tag{2. 7}$$

Further, we shall adopt

Definition 2. The entropy function multiplied by an arbitrary constant factor gives the same sensation as does the original one.

It is a plausible definition, if we regard the multiplication of the entropy function of a stimulus by a constant as a repetition of the same kind of stimulus as arbitrary number of times. We have

Lemma 2. The p, q and the ratio $m:n$ in (2. 6), and these alone, are effective to represent the elementary sensation.

This is obvious from Definition 2.

When the total numbers of degrees of freedom M is arbitrary, the sensation is given by (2. 6). For a fixed M, we have

$$\Delta M=0, \qquad \Delta P=-\Delta Q$$

and hence from (2. 7)

$$m=n.$$

The sensation is then given by

$$S'=\kappa m(\log p'-\log q'), \tag{2. 8}$$

where p and q are replaced by p' and q' in order to distinguish them from a more general case of (2. 6). If we do not distinguish (2. 6) from (2. 8), they represent the same sensation by Definition 2, so that we have the relations

$$p'=p^m, \qquad q'=q^n. \tag{2. 9}$$

On the other hand, we have

$$p'+q'=1,$$

since p' and q' stand for p and q. Thus, we have

Lemma 3. Under the restriction

$$p^m+q^n=1 \tag{2. 10}$$

for arbitrary numbers m and n, we need not distinguish the entropy function for an arbitrary M from that for fixed M, where

$$p=\frac{P}{M}, \quad q=\frac{Q}{M} \text{ and } M=P+Q.$$

The relation (2. 10) is called the equivalent primary compensation of the elementary sensation.

So far we have been concerned with the monostimulus sensation. The general stimuli can be represented by an ensemble of monostimuli. For example, the entropy function of one of them is represented by

$$H=M\log M-\sum_k P_k \log P_k$$

with the first compensation relation

$$M=\sum_k P_k$$

and its increment by

$$S=\Delta M\log M-\sum_k \Delta P_k \log P_k$$

with the second compensation relation

$$\Delta M = \sum_k \Delta P_k .$$

Among them there may be such a one as has the same sensation structure as a monostimulus, not to mention the first compensation relation.

Definition 3. Any sensation, the statistical effect of which is not distinguished from that of monostimulus sensation for a large number of stimuli, is called elementary.

From Lemma 1 and Definition 3, we have

Lemma 4. The statistically elementary sensation has a tristimulus structure.

For continuous stimuli, we have

Lemma 5. The effective elementary sensation of continuous stimuli has also a tristimulus structure.

corresponding to Lemma 4 by the well-known sampling theorem[6].

In phonetics, we recognize three formants in vowel structure. The P, Q, M in the foregoing sense are then approximately treated as the formant frequencies of vowel. Thus, from Lemma 5 we obtain the following theorem:

Theorem. The vowel has the sensation of the structure such as

$$S = \kappa(l_1 \log \omega_1 + l_2 \log \omega_2 + l_3 \log \omega_3)\,, \tag{2.11}$$

where κ is an arbitrary constant, ω_1, ω_2, ω_3 formant frequencies, and l_1, l_2, l_3 subject to

$$l_1 + l_2 + l_3 = 0 \tag{2.12}$$

and

$$\left(\frac{\omega_1}{\omega_3}\right)^{l_1} + \left(\frac{\omega_2}{\omega_3}\right)^{l_2} = 1\,. \tag{2.13}$$

It should be noted that the frequencies ω_1, ω_2, ω_3 are not necessarily subject to the first compensation relation, now. However, it is not objectionable to assume

$$\omega_1 \leqq \omega_2 \leqq \omega_3\,.$$

3. Phonetical Vowel Triangle

In the previous chapter we have seen that the sensation of the elementary vowel is defined by (2.11) with subsidiary conditions (2.12) and (2.13).

By putting

$$x = \frac{\kappa}{S} \log \frac{\omega_1}{\omega_3}\,, \qquad y = \frac{\kappa}{S} \log \frac{\omega_2}{\omega_3}\,,$$

we have

$$l_1 x + l_2 y = 1 \tag{3.1}$$

instead of (2.11), where (2.12) is taken into account. Or, by an affine transformation

$$x' = x - y , \qquad y' = y ,$$

i. e.

$$x' = \frac{\kappa}{S} \log \frac{\omega_1}{\omega_2} , \qquad y' = \frac{\kappa}{S} \log \frac{\omega_2}{\omega_3} , \tag{3.2}$$

we have

$$l_1 x' + (l_1 + l_2)\, y' = 1 \tag{3.3}$$

instead of (3. 1). Obviously $x'=0$ and $y'=0$ give the absolute boundaries of the logarithmic vowel triangle[7].

Therefore, it is theoretically anticipated that vowels are located on the lines (3. 1) or (3. 3) corresponding to the plausible sets of the indices l_1 and l_2, so that they are divided into groups corresponding to the different sets of those indices. In other words, we have

Corollary. A classification of vowels is feasible according to the possible sets of indices l_1 and l_2.

The indices l_1, l_2 and l_3 restrict the corresponding formant frequencies ω_1, ω_2, ω_3 in the argument of logarithmic function of the vowel sensation defined by (2. 11), or by

$$S = \kappa \log (\omega_1)^{l_1} (\omega_2)^{l_2} (\omega_3)^{l_3} .$$

On the other hand, we can only distinguish

$$\sin \omega t \quad \text{from} \quad \cos \omega t$$

or

$$e^{i\omega t} \quad \text{from} \quad e^{-i\omega t}$$

for an arbitrary wave with frequency ω, where t is the time variable and $i=\sqrt{-1}$. There are some plausible experiments verifying that the sensation of tones is affected by phase difference[8]. If this is the case for vowel sensation, the indices l_1, l_2 and l_3 take only the restricted values of

1 and 2

if not zero, among the integral values, by which the lower threshold of perceptible frequency difference are multiplied into l_1 or l_2.

It should be remarked that the foregoing formulation is also applicable to the vowellike consonants such as nasals and liquids, both of which have similar formant patterns to those of vowels. It sometimes happens that a formant region, i. e. one or several formants within the area, is observed in some of nasal and/or liquid spectrums. However, it is regarded as a narrow formant or its equivalent in our study.

4. Empirical Frequency Analysis of Nasals and Liquids

The formant frequencies of nasals and liquids are taken from Fant's experiments[5]. Swedish nasals /n/, /rn/, /ng/, /m/ and liquids /r/, /l/, /rl/ are pronounced by a male phonetician, whose speech is representative of standard Swedish. The

Table 1. Formant data of nasals and liquids with the parameters a' and b'

phoneme		phonetic symbol	ω_1	ω_2	ω_3	a'	b'
nasal	$/n/_1$	[n]	250 c/s	1375 c/s	2250 c/s	7.4	2.1
	$/n/_2$	[ṇ]	250	1250	2000	7.0	2.0
	/r͜n/	[ɳ]	250	1125	2000	6.5	2.5
	/n͜g/	[ŋ]	250	875	1625	5.4	2.7
	/m/	[m]	250	1000	2250	6.0	3.5
liquid	$/r/_1^*$	[r]	500	1125	2250	3.5	3.0
	$/r/_2^{**}$	[ɹ]	375	1250	1875	5.2	1.8
	$/r/_4^{***}$	[ʁ]	500	1250	2375	4.0	2.8
	/l/	[l]	250	1250	2750	7.0	3.4
	/r͜l/	[*l*]	375	1250	2625	5.2	3.2

* r_1 = rolled, apical r
** r_2 = voiced, continuant, apical r
*** r_4 = voiced, continuant back-tongue r

so-called "dark l" is omitted here, since it is difficult to identify the second formant, which is very close to the first. These data are summarized in Table 1 with the parameters a' and b', defined by

$$a' = 10 \log_{10} \frac{\omega_2}{\omega_1}, \qquad b' = 10 \log_{10} \frac{\omega_3}{\omega_2}.$$

Parameters

$$a = \frac{a' - b'}{3}, \qquad b = \frac{a' + 2b'}{3}$$

are adopted to represent vowels on the two-dimensional phonetical vowel plane[7]. Its abolute boundaries are given by

$$a' = 0 \text{ and } b' = 0 .$$

The classification of the phonemes in Table 1 are carried out for the values

$$\frac{1}{2}, \quad 1, \quad 2$$

of l_1 and l_2, since the indices in the structure of three formant vowel and/or vowel-like consonants can take only restricted values[9].

Fig. 1 shows the distribution of nasals and liquids of Table 1, as well as Swedish normal and modified vowels, the formant data of which are also looked up from Fant's analysis[5], on the logarithmic vowel plane, i. e. on the $a' - b'$ plane. The equivalent compensation relation (2. 10) is also depicted there with broken lines for the definite values of l_1 and l_2.

As regards the distribution of nasals and liquids in Fig. 1, it should be remarked that for nasals

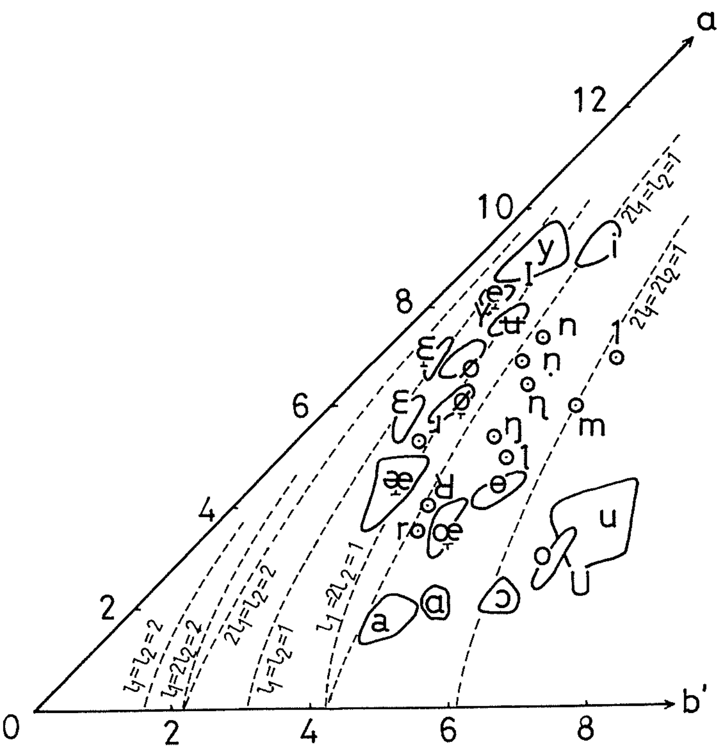

Fig. 1. Distribution of vowels, nasals and liquids on the phonetical vowel plane.

(1) the nasal cloud corresponding to [m] is located on the line $2l_1=2l_2=1$,

(2) those corresponding to [n] and [ŋ̣] is located on the line $2l_1=l_2=1$, and those corresponding to [ɳ] and [ŋ] are located between the lines $2l_1=l_2=1$ and $2l_2=2l_1=1$ nearer to the former,

and that for liquids

(3) the liquid clouds corresponding to [*l*] and [l] are located near to the line $2l_1=2l_2=1$,

(4) those corresponding to [r], [ʁ] are located on the line $2l_1=l_2=1$, while that corresponding to [ɹ] falls on the line $l_1=2l_2=1$.

By the sets of indices l_1 and l_2 of the equivalent compensation relation on which nasals and liquids as well as vowels are located in Fig. 1, we can represent the nasal and liquid sensation in the form of (3. 1) or (3. 3). Fig. 2 shows their assignment on the $x'-y'$ plane, where the values of l_1 and l_2 are also given. The neutral vowel [ɚ], which is located on the line $2l_1=l_2=2$, is supplemented to Fig. 2, aquired from the previous analysis[4] for reference. From this we can see that nasals and liquids are also divided into several groups by the ensemble of the lines given by (3. 3).

Summing up, we have the following refinement Table 2 of the natural classification of vowels, nasals and liquids.

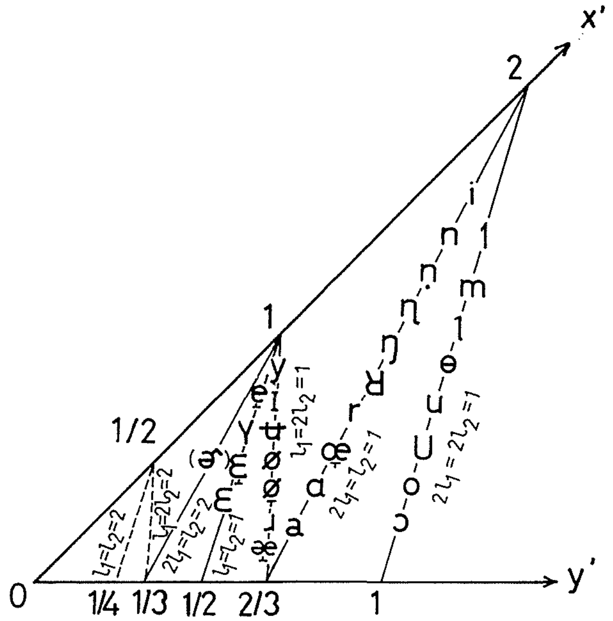

Fig. 2. Schematic classification of vowels, nasals and liquids.

Table 2. Classification of vowels, nasals and liquids by the indices l_1 and l_2

l_1	l_2	empirical phoneme
1/2	1/2	o, ɔ, u, U, θ; m, l, *l*
1/2	1	*a*, a, œ̞, i; n, ṇ, ɳ, ŋ, r, ʁ
1	1/2	æ̝, ø, ø̞, ü, y, I; ɹ
1	1	ɛ, ɛ̝, e˪, Y, (y)
1	2	(ɚ)
2	1	
2	2	

We can thus classify Swedish vowels nasals and liquids by the indices of the second compensation relation based on the information-theoretical point of view. However, the empirical verification of the phonemes, which would locate on the lines

$$l_1 = 2l_2 = 2 \text{ and } l_1 = l_2 = 2\,,$$

on which liquids are assumed to be located[3)], is still left unsettled. Although further improvements may be made in future, we may be justified to conclude that the nasals and liquids are divided into several groups respectively:

The nasal cloud corresponding to [m] falls on the line $2l_1 = 2l_2 = 1$, while those corresponding to [n], [ṇ], [ɳ] and [ŋ] on the line $2l_1 = l_2 = 1$, and

the liquid clouds corresponding to [*l*], [l] falls on the line $2l_1=2l_2=1$, those corresponding to [r], [ʀ] on the line $2l_1=l_2=1$, and that corresponding to [ɹ] on the line $l_1=2l_2=1$.

References

1) Kondo, K.: RAAG Research Notes, Third Series, No. 63 (1963).
2) Kondo, K. and Shimbo, M.: ibid., No. 67 (1963).
3) Kondo, K. and Shimbo, M.: RAAG Memoirs, 4 (1968), p. 467.
4) Shimbo, M. and Taguri, M.: RAAG Research Notes, Third Series, No. 152 (1970).
5) Fant, G.: Speech Sounds and Features (1973), pp. 36, 37, 62, The MIT Press.
6) Shannon, C. E. and Weaver, W.: The Mathematical Theory of Communication (1949), Univ. of Ill. Press.
7) Iri, M.: RAAG Memoirs, 3 (1962), p. 551.
8) Isobe, T. and Tachi, A.: Summaries of the Eighth Regular Session of the Society of Instrument and Control Engineerings (1969), p. 369.
9) Shimbo, M.: RAAG Reseach Notes, Third Series, No. 163 (1970).