HOW TO NORMALIZE YOUR OWN VOWEL SYSTEM

by

J.S.C. van Dijk

INTRODUCTION

One of the methods used in descriptive phonetics of representing vowels is by n-tuples consisting of formant frequencies. These frequencies are to be found in the speech signal by certain measuring procedures. In order to show differences between vowels within a special vowel system or to compare characteristics of different vowel systems, simple graphs in the plane (F_1, F_2 - plane) or the space (F_1, F_2, F_3 - space) are often used. In this article a procedure is proposed to normalize one's own vowel system. This in order that after normalization relative behaviour of properties within the system can be compared with corresponding properties within the normalized systems of one's fellow-creatures.

THE NORMALIZED VOWEL SYSTEM

Let S represent a vowel system consisting of N different vowels

$$(1) \qquad S = \{ V^i \}_{i=1}^{N} \quad .$$

Each vowel in S is characterized by an n-tuple consisting of n ordered numbers. These numbers are formant frequencies. Therefore we write

$$(2) \qquad V^i = (F_1^i, F_2^i, \ldots\ldots\ldots\ldots, F_n^i) \quad ,$$

with

$$F_1^i < F_2^i < \ldots\ldots < F_n^i \qquad , \quad i = 1, 2, \ldots\ldots , N.$$

Select from the set S the r-th element, $1 \leq r \leq N$. We shall call the specimen V^r the reference vocal. In agreement with (2) the reference reads

$$(3) \qquad V^r = (F_1^r, F_2^r, \ldots\ldots\ldots , F_n^r) ,$$

with

$$F_1^r < F_2^r < \ldots\ldots\ldots < F_n^r .$$

Next we join to each vowel V^i, $i = 1, 2, \ldots\ldots , N$ a new n-tuple consisting of n ordered numbers: the so-called vowel indices. These numbers are defined by

$$(4) \qquad I_j^i = \frac{F_j^i - F_j^r}{F_j^r} , \; j = 1, 2, \ldots\ldots , n.$$

In terms of vowel indices the vowel V^i reads

$$(5) \qquad V^i = (I_1^i, I_2^i, \ldots\ldots , I_n^i) , \quad i = 1, 2, \ldots\ldots , N.$$

We shall call the vowel system (1) normalized with respect to the reference (3) if each element of (1) is written in the shape (5).

THE REFERENCE VOCAL V^r

Estimates for formant frequencies of vowels can be obtained using appropriate chosen models of the vocal tract (Fant 1960; Mol 1970; Bonder 1979). The simplest vocal tract model which represents a vocal in geometrical characteristics is the cylindrical tube, closed at one end. Natural frequencies of that tube are

$$(6) \qquad F_m = (2m - 1) \frac{c}{4\ell} , \quad m = 1, 2, \ldots\ldots ,$$

in which c denotes the velocity of sound propagation in the tube and ℓ the length of the model. The first n natural frequencies of (6) are estimates for the first n formant frequencies of this 'straight tube' vowel. We shall select this vowel as the reference vocal V^r. So, (3) takes the shape

(7) $$V^r = \left(\frac{c}{4\ell},\ 3\frac{c}{4\ell},\ \ldots\ldots\ldots\ldots,\ (2n-1)\frac{c}{4\ell} \right),$$

so that the vowel indices (4) of the i-th vowel read

(8) $$I^i_j = \frac{F^i_j}{F^r_j} - 1 \quad \text{with} \quad F^r_j = (2j-1)\frac{c}{4\ell}\ ,\ j = 1, 2, \ldots\ldots, n.$$

In practice a reference of the kind (7) is not at hand. Therefore we shall determine (7)in a constructive manner. This can be accomplished as follows. In terms of vowel indices the distance of the i-th vowel of the system to the reference (7) is

(9) $$D^i = \left\{ \sum_{i=1}^{n} (I^i_j)^2 \right\}^{\frac{1}{2}},\quad i = 1, 2, \ldots\ldots, N.$$

Let us define the norm D of the system (1) as the root mean square value of the N distances in (9). Then we have

(10) $$D = \left\{ \frac{1}{N} \sum_{i=1}^{N} (D^i)^2 \right\}^{\frac{1}{2}}$$

From (8), (9) and (10) follows that D^2 is quadratic in ℓ. In consequence of this it is easy to minimize D. We shall select as the reference vowel that 'straight-tube' vowel with tube length ℓ for which (10) is minimal. A simple calculation shows that in this case

(11) $$F^r_1 = \frac{c}{4\ell} = \frac{\displaystyle\sum_{i=1}^{N}\sum_{j=1}^{n}\left(\frac{F^i_j}{(2j-1)}\right)^2}{\displaystyle\sum_{i=1}^{N}\sum_{j=1}^{n}\frac{F^i_j}{2j-1}}\ .$$

(11) determines the reference vowel (7) completely. Addition of this artificial reference vowel to the system (1) makes it possible to normalize the system S according to the proposed procedure.

EXAMPLE

The second and third column of table 1 comprise the first two measured formant frequencies of the vowel system of a trained male Dutch speaker (Koopmans-van Beinum, 1980). The vowels originate from monosyllabic words. From these data the reference vowel (7) has been calculated by application of (11). Using (8) the first two vowel indices have been determined with (8). In order to make a graph in the I_1, I_2 - plane unit lengths along the separate axes have been chosen. These lengths are denoted with d_1 and d_2, respectively, and represent the standard deviations in the I_1 and I_2 direction. The values of I_1 in units d_1 and I_2 in units d_2 are given in the third and fourth column. Fig. 1 is the graph of the normalized vowel system.

	F_1(Hz)	F_2(Hz)	$I_1(d_1)$	$I_2(d_2)$
u	308	809	-1.39	-1.51
o	407	817	-0.78	-1.51
ɔ	506	959	-0.19	-1.23
ɑ	657	1266	+0.71	-0.65
a	717	1437	+1.10	-0.34
y	296	1861	-1.45	+0.49
ø	372	1614	-1.00	+0.01
œ	451	1446	-0.52	-0.31
i	292	2470	-1.50	+1.63
I	385	2156	-1.07	+1.05
e	392	2129	-0.87	+1.11
ε	605	1741	+0.42	+0.25

$F_1^r = 537$ Hz

$d_1 = 0.31$

$d_2 = 0.32$

Table 1

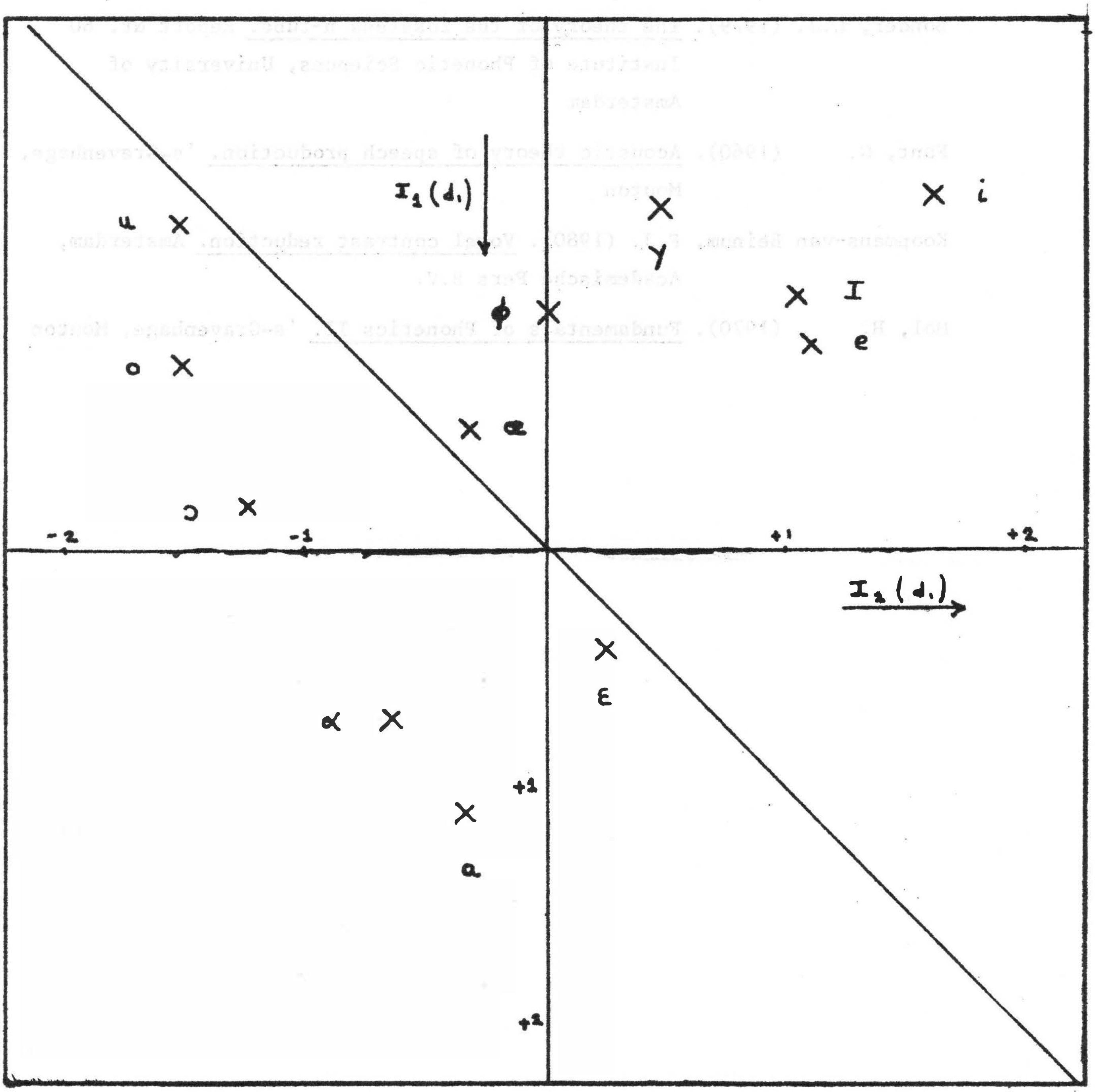

Fig. 1

REFERENCES

Bonder, L.J. (1979). *The theory of the lossless n-tube.* Report Nr. 60 Institute of Phonetic Sciences, University of Amsterdam

Fant, G. (1960). *Acoustic theory of speech production.* 's-Gravenhage, Mouton

Koopmans-van Beinum, F.J. (1980). *Vowel contrast reduction.* Amsterdam, Academische Pers B.V.

Mol, H. (1970). *Fundamentals of Phonetics II.* 's-Gravenhage, Mouton