A MODEL OF THE VOWEL SYSTEM OF DUTCH

H. MOL*

At the spring-meeting of the Acoustical Society of America at Washington in 1965 I discussed the backgrounds of the twin-tube model of the vocal tract. I pointed out how to transform the well-known formant-formula

$$\tan \frac{\omega l_1}{c} \tan \frac{\omega l_2}{c} = \frac{S_1}{S_2} = k$$

into

$$\cos \frac{\omega l_a}{c} = \frac{1-k}{1+k} \cos 2\Delta \frac{\omega}{c}$$

by introducing a new coordinate, the *eccentricity* Δ that describes the position of the joint as its distance to the middle of the twin-tube.

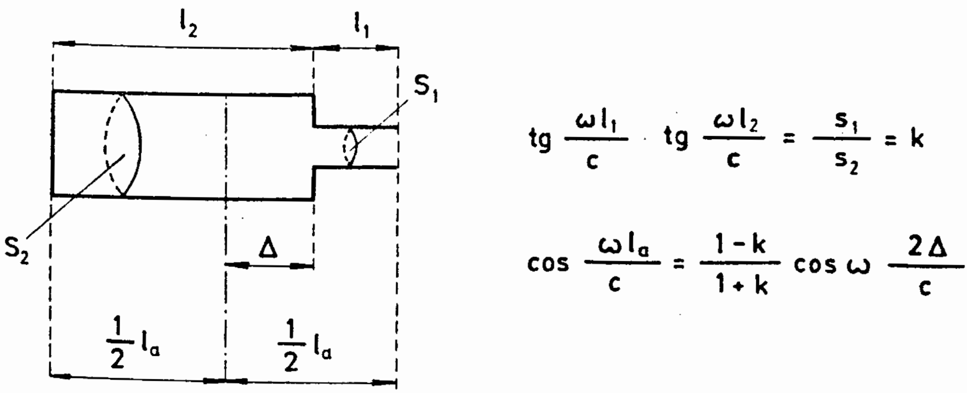

Fig. 1.

The cosine formula can be solved either graphically or by means of a computer. In describing the vowels of Dutch we used both methods: the simple graphical method in order to get a quick insight into the nature of the vowel diagram and the computer method for more sophisticated applications.

* Institute of Phonetic Sciences of the University of Amsterdam.

The geometry of a twin-tube is completely determined by its parameters l_a, Δ and k. The 12 vowels of Dutch in isolated key words, carefully pronounced by a native speaker, can be described by a set of 12 twin-tubes producing the same formants. The parameters of each twin-tube can be chosen on basis of the study of X-ray photo-

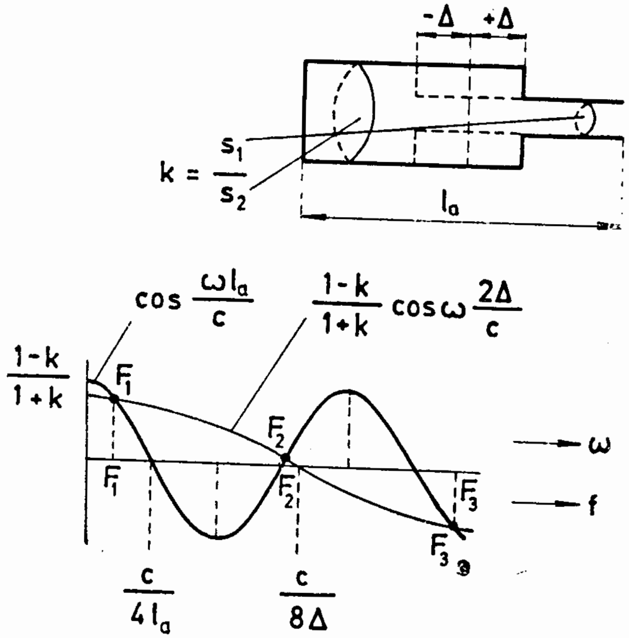

Fig. 2.

graphs, palatograms and formant-measurements of that speaker. By now it is well-known that the absolute formant positions — hence the absolute values of the parameters—do not and indeed cannot function phonemically. Instead, it is the contrasts between the formants of the vowels that form a system. To put things more precisely, though the *absolute* formant positions may vary between speakers, the *relative* positions of the vowels in the complete vowel diagram in the F_1F_2 plane are always the same for native speakers of the same language, provided these speakers have no apparent speech defects and are free from outspoken dialectical peculiarities. The formant positions depicted in the slide refer to a talker with about average anatomic dimensions. We may pin the phonemic system on him as well as on any other normal speaker of Dutch.

The acoustical length of the vocal tract of a speaker depends on the vowel he pronounces in an isolated word. For instance, in [*i*] the vocal tract is short, because, acoustically, it begins somewhere behind the teeth, the larynx is usually somewhat lifted, whereas the back of the tongue is low. On the other hand, in [*u*] the lips are protruded, the larynx is in a low position, whereas the back of the tongue is high. In that case the axial acoustical length of the vocal tract is at its maximum. In [*a*]

the length is about average. It suffices to confine ourselves to 3 different contrasting lengths of the vocal tract in Dutch for depicting the phonemic system, but this is not a decisive factor as long as the contrasts are being conserved. As to the parameter

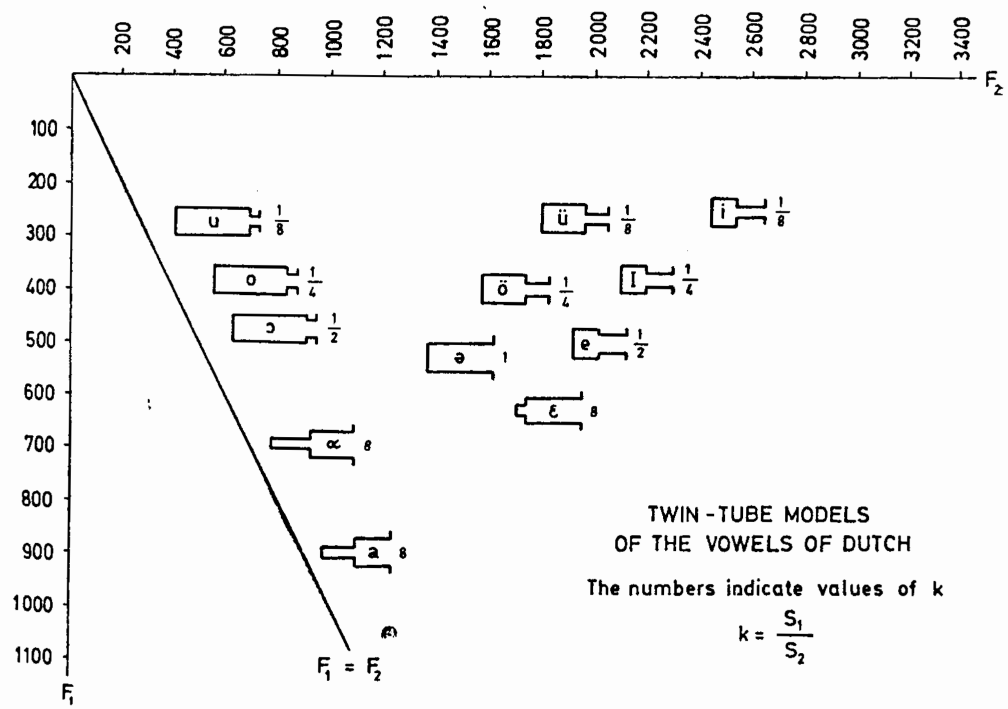

Fig. 3.

$k = \dfrac{S1}{S2}$, five suitably chosen values will do. For the eccentricity Δ, three contrasting values suffice.

Strictly speaking, the twin-tube model is only a mathematical concept involving several idealisations. It is very tempting, and also very rewarding, to *materialize* the twin-tube model in the shape of two hard-plastic tubes, driven by a suitable sound source imitating the real larynx. For this artificial larynx we have chosen an acoustical siren, visible on slide 4. Compressed air is being let in through a triangular opening in series with a rotating slit. In that way a continuous series of air puffs, having the necessary shape, is generated. Each air puff begins gradually but ends with an abrupt cut-off which generates powerful damped oscillations in the artificial vocal tract. The thusgenerated vowels do not sound any sillier than the sustained vowels produced by life speakers.

In Dutch there are sequences of two vowels that are traditionally called diphthongs by many, though they are nothing but vowel clusters in which both elements carry phonemic information. We shall not deal with these diphthongs here.

Moreover, there are vowel clusters in which only the first element carries phonemic information, whereas the second element is merely an audible transition without a phonemic function. This type is often called a diphthongal vowel. We shall not discuss it here.

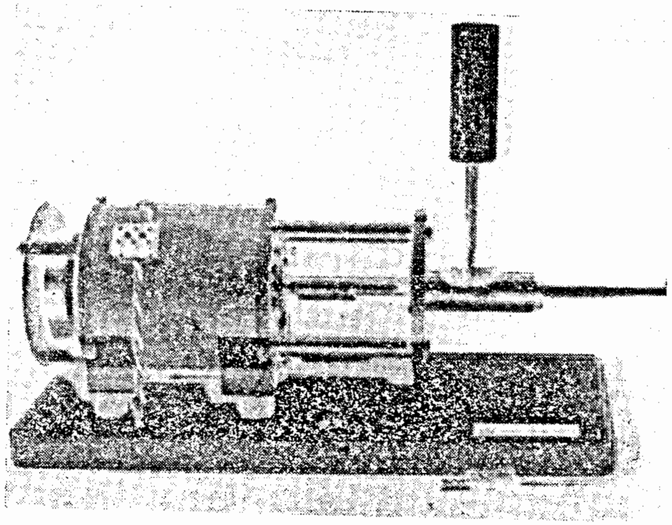

Fig. 4.

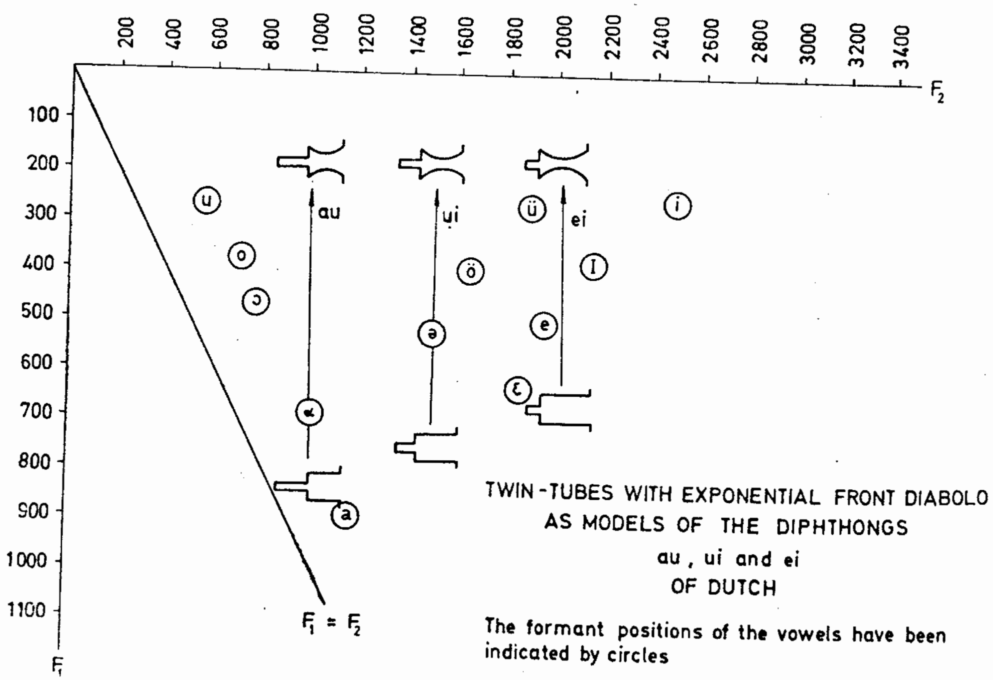

Fig. 5.

In Dutch, however, there is a special class of diphthongs of a *dynamic* nature, namely the speech sounds [*au*], [*ui*] and [*ei*]. They consist of two vowel-like elements having the same (or slightly different) F_2, gliding into each other by a rapid shift of F_1.

It is very probable that this shift, together with the position of F_2, carries the phonemic cue. Especially towards its end, however, this shift is so rapid that it seems to fall out of step with the comparatively slow articulatory movements. For the description of these three diphthongs we want a model that produces a practically constant F_2 and an ever-accelerating glide of F_1. Such a dynamic model can be derived from the twin-tube model by squeezing its mouth-tube into the shape of a diabolo. For reasons of mathematical simplicity we suppose that the diabolo consists of two identical exponential horns. In that case squeezing can be reduced to simply changing the flare of the horn.

The time allotted to this paper does not allow me to present the convincing calculations but they will be published elsewhere.