Proc. 5th int. Congr. phon. Sci., Münster 1964, pp. 422–425
(S. Karger, Basel/New York 1965).

Some Problems around the Growth of the Vocal Tract

By H. Mol, Oegstgeest

In phonemic theory the anatomic differences between the articulatory organs of men, women and children have never been taken seriously into account. The same can be said of the differences within one of these three groups. The level of abstraction on which the phoneme is defined is so high above, or in any case so distant from the actual mechanisms involved in the process of speaking and listening that the indifference of the phonemicists towards these anatomic variations can be easily imagined: the underlying doctrine of phonemics is the statement that isolated words are 100% intelligible, no matter what.

As long as the acoustic theory of the vocal tract and reliable measurements of the vowel formants were still in the cradle much faith could be attached to compensatory mechanisms that were supposed to alleviate many of the anatomic differences between the speakers of the same language.

In the light of modern phonetics two questions can be asked:

1. Are compensatory mechanisms powerful enough to alleviate the acoustic results of the gross anatomic differences between men, women and children?

2. If they are, are they actually brought into play by the speakers? Both questions must be answered in the negative.

By applying, for instance, *Webster's* horn equation it can be proved that the frequencies of *all* formants of the vocal tract are inversely proportional to the length of that tract as measured between the lips and the vocal cords. From this fact a very interesting conclusion may be directly drawn.

Suppose a young speaker has mastered to produce the inventory of vowels characteristic of his language and, during the growth of

his vocal tract, he does nothing to compensate for the acoustic results of this axial growth, all his 'adult' formants will be lower by the same factor than the corresponding 'juvenile' formants. It means that he then keeps on giving the same commands to his articulators. Measurements prove that this situation is the normal.

Though the authors themselves do not stress the point, this proporti nality factor can indeed be found in the well-known formant measurements of *Peterson* and *Barney*[1], who presented separate averages for the formants of men, women and children in carefully pronounced words containing one vowel in medial position.

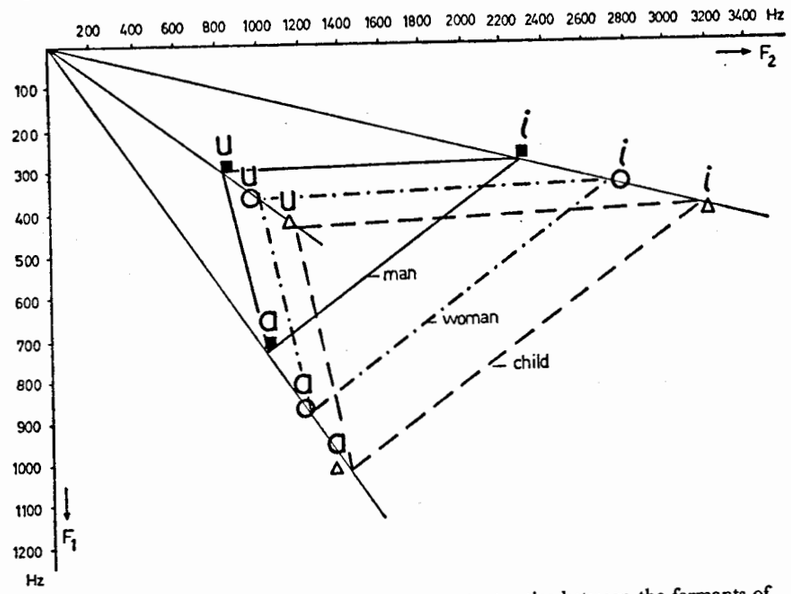

Fig. 1. Illustration of the fact that there are constant ratios between the formants of men, women and children.

In Fig. 1, based on the measurements of *Peterson* and *Barney*, for the sake of clarity, only the three vowels forming the familiar vowel triangle of *Hellwag* have been brought to the fore, though the properties of proportionality also hold for the other vowels. Within reasonable limits of accuracy the triangles can be derived from each other by simple multiplication from the zero point of the frequency scales. The corresponding vowels of men, women and children are

on straight lines through that zero point, the so-called origin, provided linear scales are chosen for both formants. The ratio between the formants of women and men equals 1.16. The same factor relates children and women, whereas between the formants of children and men the ratio 1.35 obtains.

While learning to speak the child does not (and indeed cannot) imitate the formants of the adult speakers in its environment. Unlike a parrot, it does not give exact imitations on an acoustic basis. In the long run the child learns to master the muscular activity necessary for producing the *same number* of perceptively different vowel sounds used for coding purposes by adult speakers of the same language in situations where the need for utmost clarity is greatest. Gradually the child learns to choose the, in his language, normalized number of perceptually different tongue positions realizable within the limiting boundaries of the teeth, the back, the roof and the floor of the mouth. He uses and needs his ear in this process but certainly does not adapt the absolute values of the formants to the example.

Once settled in a certain period of life the neurological programmes for controlling the articulatory muscles do not change materially as the child grows up. Apparently the tongue keeps the same relative position in the vocal tract for each vowel. Lengthening of the vocal tract makes the formants shift downwards and there are no compensatory attempts on the part of the speaker to make them stay at their 'juvenile' positions. Obviously the role of the often postulated auditory feed-back mechanism for regulating the articulatory movements is an extremely minor one.

Every speaker has his own vowel system. The range of the frequencies of his formants is influenced by the dimensions of his vocal tract. The listener has to cope with the difficulty that the absolute values of the formants of the vowels vary from speaker to speaker[2]. He is helped, among other things, by the fact that in a given language all speakers have the same *number* of vowels. A listener is able to adapt himself in a surprisingly short time to the formant positions of a speaker. When he is not able to do so, for instance when isolated words of unknown speakers are presented to him, his identifications of the intended vowels become inaccurate[1].

If we admit that the exact acoustic imitation of the vowels is not the aim of learning to speak, we can understand why the inherited shapes of parts of the vocal tract, like the hard palate, can withstand

'compensatory' actions. It is a well-known fact that quite often the voice of a speaker resembles that of his father as soon as his larynx and his vocal tract have reached their final dimensions.

In general it can be said that the encouragement for an articulatory programme to 'settle' is not the result of successfully matching an example. This encouragement must stem from the dawning of the idea that the sounds produced really function. The parrot is in a different position: he is a real imitator. We have established the fact that a parrot learns to speak without training. He is able to store a word or a sentence heard only once at a certain time, often under conditions of psychological stress, in his memory and to reproduce this utterance suddenly at a considerably later moment. Apparently he possesses a built-in inventory of acoustically 'labelled' articulatory programmes; as soon as he hears a sound he immediately selects the corresponding articulatory programme. Let us call this gift 'parroting'. Every now and then one hears about children who seem to have the gift of parroting. They are very silent, so long that the parents become uneasy about it, until suddenly they produce 'complete' words.

Besides permanent growth the vocal tract is subject to temporary length-variations controlled by the will of the speaker.

The axis of the vocal tract can be lengthened by moving the larynx downwards, moving the hump of the tongue upwards and backwards and by pouting the lips. There even are formant shifts that can only be explained by a length-variation of the vocal tract, for instance the shift from [a] to [ɑ] in Dutch.

References

1. *Peterson, G.E.* and *Barney, H.L.:* Control methods used in a study of the vowels, JASA Volume 24, Number 2, p. 175–184, March 1952.
2. *Ladefoged, P.* and *Broadbent, D.E.:* Information conveyed by vowels, JASA 29, p. 98–104 (1957).

Author's address: H. Mol, Prinses Beatrixlaan 28, *Oegstgeest* (Netherlands).

Discussion

Sovijärvi (Helsinki): Herr *Mol* hat darauf hingewiesen, daß die Relation des ersten bzw. zweiten Vokalformanten bei den Männern und Kindern 1:1,35 sein würde (vgl. Fig. 1). Nach meinen Messungen, die ich von der laufenden finnischen Rede eines Mannes und eines 10jährigen Jungen gemacht habe, schwanken die Formantgipfelhöhen des F1 in demselben Bereich, welche Tatsache dadurch erklärt werden kann, daß jeweils die phonematisch am besten passenden F1-Höhen auf sehr verschiedene Weise wegen der wechselnden Grundtonhöhen erreicht werden. Dagegen ist die erwähnte Relation 1:1,35 betreffs der F2-Höhen auch nach meinem Material stichhaltig (vgl. meinen Artikel in «Phonetica, Symposium Trubetzkoy», 1958).