https://doi.org/10.3176/lu.1969.2.07

ARVO EEK, MART REMMEL (Tallinn)

SOME REMARKS CONCERNING SPEECH PRODUCTION

0. J. S. Perkell's monograph [1] is one of the first works where, alongside with hypotheses underlying the "physiological model" of speech production, measurements of the vocal tract which in their turn serve as a basis for the hypotheses have been comprehensively presented. The value of such treatments for phonetics is obvious (among other things in the aspect of cross-language study). The main results of J. S. Perkell's work are both novel and noteworthy. They have, however, also evoked the few notes presented below.

1. 1. The only conceivable and logical requirement to be considered when choosing a reference system for lateral cinefluorograms is that the system should f o l l o w the vocal tract (especially with a view to transforming the lateral measurements into cross-sectional areas and, in the final analysis, to calculating the spectra) and/or the behaviour of articulators (especially with a view to describing the dynamics of articulation).

Of the harmonic contours following the curvature of the vocal tract (in the region of the hard and the soft palate), the arc of an ellipse would probably be completely satisfactory. Practically it is possible to substitute the arc of a circle for the arc of an ellipse; measurement and calculations are relatively easier to perform in case of the arc of a circle. Further, it is practically possible to replace the arc of a tangent circle by the arc of a circle passing through three relatively immovable points sufficiently visible on X-ray frames: the tip of the crown of the anterior maxillary incisor, the posterior nasal spine, the anterior lower edge of the second cervical vertebra. This is possible because these three points are at approximately equal distances from the outer boundary of the vocal tract. In this way a circle is obtained whose centre lies sufficiently close to the centre of the tangent circle of the vocal tract. This circumstance has been utilized by J. M. Heinz and K. N. Stevens.[2]

In the rest their system follows the outer boundary of the vocal tract so that the coordinates would be as perpendicular to it as possible. The following of the outer boundary of the vocal tract is correct in the region of the front teeth, alveoli and pre-palate [3], but not in the pharyngeal and laryngeal regions. A system which should be more economical for the latter regions is presented below.

---

[1] J. S. P e r k e l l, Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study, Cambridge, Massachusetts 1968.

[2] J. M. H e i n z, K. N. S t e v e n s, On the Relations Between Lateral Cineradiographs, Area Functions, and Acoustic Spectra of Speech. — Paper A44 in: 5e congrès international d'acoustique. Rapports conférences particulières 1a; Liége 1965.

[3] For measurements and calculations it is essential that the inner boundary of the vocal tract should also behave as economically as possible with regard to the reference system. The coordinates which are perpendicular to the outer boundary of the vocal tract between the maxillary front teeth and the pre-palate are generally suitable for the corresponding parts of the tongue as well.

1.2. It is also possible to follow only the movements of certain parts of the vocal tract which are more essential from the acoustic and articulatory points of view, without a system involving the whole vocal tract. Such is also J. S. Perkell's approach. But the whole system of measurements as presented by him is to a certain extent ambiguous (fortunately the ambiguities have no essential influence on J. S. Perkell's conclusions). It would be ideal to fix definite points on both the immovable structure as well as on the movable structure, the latter being measured with reference to the former (following, for example, Perkell's method who uses lead pellets fixed to the tongue surface, provided that they do not interfere with articulation).

It seems that for greater accuracy it is useful and economical to include at least some part of the measurements describing the movements of articulators in a strict system comprising the whole vocal tract.

1.3. When there are more informants than one, a normalizing system is needed, i. e. similarly marked coordinates for different informants should describe one and the same anatomical structure. One such normalizing system will be described below. For the purpose of normalization this system makes use of some anatomically probable data on the length of parts of the vocal tract. In principle, it also corresponds relatively well to the approach of traditional phonetics, and renders possible a comparative description of

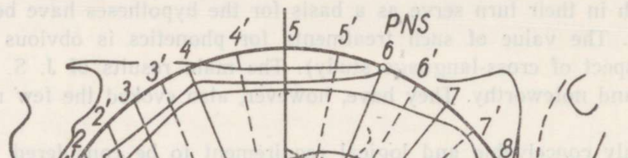

Fig. 1. X-ray tracing with landmarks and reference coordinate system.

data for different informants from the viewpoint of the place of articulation. (A more general system based on the arc of an ellipse can be normalized with the aid of the theorems of Poncelet.)

2.1. In the following we shall describe the reference system that we use.[4]

The centre of the system, *0*, is fixed at equal distances from three points of immovable structures: from the tip of the crown of the anterior maxillary incisor, *I*, from the posterior nasal spine, *PNS*, and from the anterior lower edge of the second cervical vertebra, C_2 (see Fig. 1). The radii of the circle passing through these three points

[4] The system is based in its essentials on that of J. M. Heinz and K. N. Stevens:

are used as polar coordinates. The polar coordinates begin with the radius which is the bisector of the angle *1-0-PNS* (usually it coincides with the boundary between the prepalate and the mediopalate, Fig. 1, *4*) and proceed as far as C_2 (*10*). The intermediate coordinates are obtained separately for the arcs *4—6* and *6—10* by their (repeated) bisecting. The interval step of such coordinates need not be constant throughout the system.

10 is perpendicular to *5* (and thus need not be a radius). The rectangular coordinates *10a* and *10b* are placed parallel to *10*. The length of the interval step is equal to half of the distance between the lower anterior edges of C_2 and C_3.[5]

11 through *16* (and the following, if necessary) are at a 30° angle[6] relative to *10* and reach as far as the prolongation of *5*. The length of the interval step here equals one-third of the distance between *14* (runs into the lower anterior edge of C_4) and *11* (runs into the lower anterior edge of C_2).

2 (defined by the pit in the front part of the neck of the upper incisor) is parallel to *4* and reaches as far as the prolongation of *10*. Coordinates between *2* and *4* are obtained by (repeatedly) bisecting the straight line segment between the points where these two coordinates intersect with the prolongation of *10*. For a more detailed description of dental articulation the distance between *1* (based on the tip of the upper incisor) and *2* is divided into two equal parts.

The lateral measurements of the vocal tract are obtained as distances between the points where the coordinates intersect with the outer and inner boundary of the vocal tract, i. e. on the one hand with the hard palate, the velum, the rear wall of the pharynx and on the other hand with contours of the tongue and the epiglottis. Depending on the precision of the research, besides the basic coordinates (*1*, *2*, *3*, *4*, etc.) intermediate coordinates obtained by bisecting the intervals can be used (such as *2′*, *3′*, *4′*, etc.).

Some additional coordinates are introduced in order to describe the dynamics of gestures of articulators:

L_h — the height of the lip aperture; measured as the distance between the tangents (perpendicular to *5*) to the upper and lower lip;

I_d — the distance between the tips of the anterior upper and lower incisors (describes the movements of the mandible);

M — the distance of the tip of the anterior lower incisor from the prolongation of the *10*th coordinate (describes the movements of the mandible);

U_h — the height of the uvula; measured as the distance between the tangent (perpendicular to *5*) to the upper curve of the uvula and *10*;

U_w — the distance between the back wall of the uvula and the rear wall of the pharynx; measured along the perpendicular to *5* from *7*, *7′*, *8*, *8′* or *9*;

Lar — the distance between the anterior end of the ventricle and the tip of the upper incisor (describes the movements of the larynx);

H — the distance of the body of the hyoid bone from the tip of the upper incisor (describes the movements of the hyoid bone);

T_c — the length of the tongue contact (the distance from the most anterior to the most posterior point of contact between the tongue and the outer boundary of the vocal tract).

3. J. S. Perkell's statement (4.3), "The feature tense-lax will not play a role, however, when there is no increase in pressure, as in the sonorant consonants", can hardly be correct. First, the impression of the tenseness v. laxness of consonants derives from only two pairs of consonants ([t, d] and [s, z]), and second, there do exist languages

[5] When a five-inch electron optical image intensifier tube is used in cinefluorographic processing it may happen that the lower cervical vertebrae of informants with bigger head dimensions are left outside the frame. In that case at least one static X-ray shot is taken which, after its reduction to normal size, enables us to extrapolate the reference point C_3 for the coordinate system.

[6] Depending on the aims of one's research (and on the sound types), a combination of coordinates is to be chosen which is necessary and sufficient for carrying out measurements.

where the feature tense v. lax can distinguish sonorant consonants. The following examples are from Estonian where the opposition of the so-called degrees of quantity in consonants results (as we understand it) from the simultaneity of the features syllabicity and tenseness.[7]

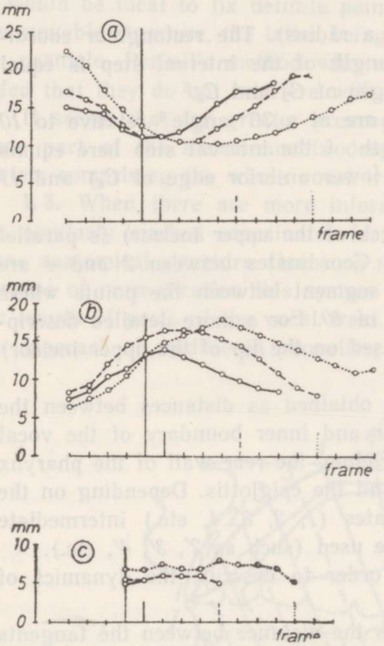

Fig. 2. Graphs of the measurements of the Estonian words *kana* (———), *kanna* (— — —), *Anna* (ill.; . . . . .). Informant O. P. (woman).

a — the distance of the dorsum of the tongue from the hard palate (coordinate *5*); *b* — the distance of the root of the tongue from the rear wall of the pharynx (coordinate *10*); *c* — the length of the alveolar contact of the tongue (T_c).

The abscissa plots time (the succession of frames). The ordinate plots measurements of cinefluorograms in millimetres. The first vertical line marks the beginning of the consonants, the next verticals mark the end of corresponding consonants.

N o t e: the graphs represent measurements from the end of the quasi-stationary segment of the vowel preceding the consonant up to the quasi-stationary segment of the vowel following the consonant. The segmentation was performed by comparing the X-ray frames with the synchronous dynamic sound spectrograms of the same material.[8]

Fig. 2 displays the measurements *5*, *10* and T_c of [n] in three degrees of quantity which we treat as /n/, the "geminate"[9] with a lax first component, and the "geminate" with a tense first component.

T_c measures the length of the tongue contact on the alveoli which increases with an increase of the so-called degree of quantity: 5, 6 and 8 mm respectively (the same tendency appears on palatograms as an increase of the area of both the alveolar and the side contact[10]). The dorsum of the tongue rises in the direction of the hard palate (*5* — the distance between the dorsum of the tongue and the hard palate) least of all in the so-called first degree of quantity ($5 = 11$ mm); and most of all in the so-called third degree of quantity ($5 = 10$ mm). The movement tendencies in the second and third degrees are similar (the tongue starts moving in the direction of the following vowel right after the culmination phase of [n] and continues to do so up to the release, whereas the [n] of the first degree of quantity reaches its maximum only immediately

[7] Naturally, this kind of approach to the opposition of the degrees of quantity in consonants is missing in the hitherto existing, mainly taxonomic phonology of Estonian, but for us it appears to be clear at least in the aspect of markedness.

[8] For the method of cinefluorographic processing, see: Г. Л и й в, А. Э э к, О проблемах экспериментального изучения динамики речеобразования: комплексная методика синхронизированного кинофлуорографирования и спектрографирования речи. (With summaries in Estonian and English.) — Eesti NSV Teaduste Akadeemia Toimetised. Bioloogia 17, 1968, pp. 84—89.

[9] The phonetic equivalent of a "geminate" here is a long sound (as contrasted with a short one). The phonetic interpretation of the "geminate" in the present case excludes the possibility of looking for a culmination phase for each of the components of a "geminate".

[10] Cf. P. A r i s t e, Eesti ühiskeele palatalisatsioonist. — Katselisfoneetilisi tähelepanekuid. Acta Universitatis Tartuensis B-L_2, Tartu 1943, pp. 26—28; S. S h i b a t a, A Study of Dynamic Palatography. — Annual Bulletin No. 2. Research Institute of Logopedics and Phoniatrics. University of Tokyo, Tokyo 1968, pp. 28—36.

before the release of the consonant). This tendency (of correlation between measurements and the degree of quantity) is even more conspicuous in the movement of the root of the tongue away from the rear wall of the pharynx (*10*: 14, 16.5, 18 mm, respectively). It seems that the changes in the measurements *5* and *10* accompany the changes in T_c (being partly a result of the differences in duration as well), i. e. a higher degree of quantity means an increase in the articulatory effort which makes the tongue move closer towards the hard palate and away from the rear wall of the pharynx. The relative analogy of gemination and tenseness deserves further attention.

It thus seems inevitable that this specific physiological formulation (of J. S. Perkell) [11] has to be abandoned (as has been done many a time in the history of fortis-lenis/tense-lax) and that account should be taken only of "greater muscular effort". At the same time this provides indirect evidence of the fundamental nature of the feature tenseness-laxness. It is needless to say that smaller sound classes may allow of more specific formulations of tenseness-laxness.

[11] J. S. Perkell, *op. cit.*, § 4.3: "The concept of tense seems to be related to greater muscular effort for the consonants as well as for the vowels. For vowels, the greater muscular effort is applied to moving the tongue body, but for consonants it seems to be related to resisting increases in intraoral pressure." See also § 3.8.