ARTICULATORY PARAMETERS

Peter Ladefoged, Phonetics Lab., Linguistics Department, UCLA, Los Angeles, CA 90024, USA

The main report for this session gives an excellent summary of recent research on speech production. I would like to try to summarize this summary by listing and discussing the articulatory parameters that need to be controlled in a model of the speech production process. Obviously this could be done at various levels of generality. For example, one could choose to model the various muscular forces acting on the tongue, as suggested by Fujimura and Kakita (1978), or one could model the results of those forces as described by Harshman et al. (1977). Similarly one could specify the gross respiratory movements as Ohala (1975) has done, or more simply the variations in subglottal pressure that result from those movements. On another dimension of generality, one could try to describe just those articulatory parameters required for a particular language, or the larger set that would produce all possible linguistic differences, or even those that would go still further and allow one to distinguish all the personal characteristics of individual speakers.

I have chosen to specify speech production in terms of the minimal set of articulatory parameters given in Table 1. They will (hopefully) account for all linguistic differences both within and between languages, but may not distinguish between speakers. There is a lot of guess-work involved in setting up a list of this kind. Some of the parameters (eg 1, 2, 8, 9, 11, 16) can be defined fairly precisely, but others (eg 5, 6, 7, 14) are less firmly established.

The parameters listed may be thought of as corresponding to what is controlled rather than to movements of anatomical structures such as the jaw or the ribcage. This is a somewhat controversial point in that Lindblom and Sundberg (1971) have proposed that it is more appropriate to model tongue movements with respect to a moving mandible, rather than simply modeling the vocal tract shapes that result from these tongue movements. But it seems to me that if one is trying to state the parameters that are used in controlling articulatory actions, then Lindblom's own work (Lindblom et al., 1978) shows that speakers may rely on a great deal of

Table 1

A necessary and sufficient set of articulatory parameters.

1. Front raising	9. Lip width
2. Back raising	10. Lip protrusion
3. Tip raising	11. Velic opening
4. Tip advancing	12. Larynx lowering
5. Pharynx width	13. Glottal aperture
6. Tongue bunching	14. Phonation tension
7. Lateral tongue contraction	15. Glottal length
8. Lip height	16. Lung volume decrement

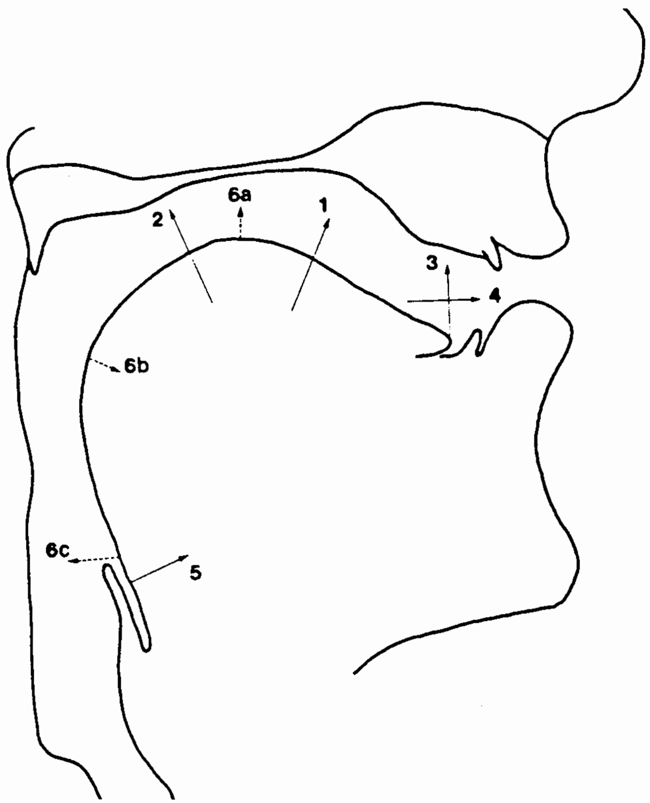

Figure 1

The movements of principal portions of the tongue associated with the first 6 parameters in Table 1.

compensation between movements of the jaw and those of the tongue. What they control are the vocal tract shapes, i.e. the relative magnitudes of the cross-sectional areas of the mouth and pharynx. The underlying parameters may therefore be as shown in Table 1.

The first six parameters are concerned with the position of the tongue relative to the roof of the mouth and the back wall of the pharynx. Most of these also involve movements of the soft palate and the pharynx, and it is only a convenient simplification to regard them as merely movements of the tongue. They are really parameters for the control of vocal tract shape.

For each of the first five parameters there is one portion of the tongue which makes the largest movement, and this portion may be used to name the parameter as a whole. These movements are shown in Figure 1.

It should be emphasized that each parameter specifies more than the movement of a single point. Thus the first parameter, front raising, specifies the degree of raising or lowering of the front of the tongue, and also the concomitant advancement or retraction of the root of the tongue. To say that a given sound has a certain degree of front raising means that the tongue as a whole may be said to be deviating from a neutral reference position to that degree. The arrow marked 1 in Figure 1 shows the potential movements of that part of the tongue that moves most with variations in front raising. Other points will move to a lesser degree.

The first two parameters, front raising and back raising (arrows 1 and 2), have been fully described in a series of recent publications (Harshman et al. 1977, Ladefoged et al. 1978, Ladefoged and Harshman 1979). These parameters enable us to give explicit formal descriptions of the movements of the tongue of an average speaker, such that we can characterize, fairly accurately, at least the non-rhotacized vowels of English.

It is obviously of interest to phoneticians to compare descriptions in terms of front raising and back raising with more traditional descriptions in terms of the highest point of the tongue, but unfortunately this cannot be done at the moment. The problem with these traditional descriptions is that no one has as yet shown how to interpret them unambiguously. Given the height and degree of backness of the highest point of the tongue (and given that all the other parameters such as pharynx width have neutral values) it is not yet known how (or even if) the position of the tongue as a whole may be described.

The remaining parameters in Figure 1 have not been investigated as fully as the first two. It seems clear that there must be two degrees of freedom to movements of the tip of the tongue, as suggested by the arrows marked 3 and 4. There are many sounds which involve advancing or retracting the tip of the tongue while raising or lowering it in varying degrees. But we do not really know exactly what it is that is controlled, nor how these two parameters are related to one another. Furthermore, as Ohala (1974a) has pointed out, these movements may also affect the back of the tongue. It is impossible to do more than guess at a full mathematical specification of these parameters.

The fifth parameter, pharynx width, has been discussed extensively by Lindau (1979). For most languages, the position of the body of the tongue in vowels can probably be described very adequately in terms of the two parameters, front raising and back raising. But there are a number of languages such as Akan and Igbo, in which the width of the pharynx is independent of the height of the body of the tongue.

The three dotted lines in Figure 1 represent an estimate of the effect of the sixth parameter, tongue bunching. This estimate is based on an analysis of only five speakers of American English saying the vowel /əɹ/ as in "heard", and should be regarded as very tentative. Line 6a indicates a bunching up of the front of the tongue, 6b a concomitant increase in the opening of the vocal tract in the upper part of the pharynx, and 6c a considerable narrowing in the lower part of the pharynx. All these co-occur in tongue bunching in American English. But it should be noted that vowels of this kind are very unusual, and are likely to occur in less than 1% of the languages of the world (Maddieson, personal communication).

The final parameter associated with adjustments of tongue shape is lateral tongue contraction, which occurs in the production of laterals. Because the tongue is an incompressible mass, decreasing the lateral dimension must cause an increase in some other dimension. But we do not know how the narrowing movement is controlled. If speakers are aiming to control vocal tract shape, then decreases in tongue width may be complemented by

movements of the tongue within the mandible, absorbing potential increases in tongue height.

In addition to movements of the tongue (and the concomitant movements of the pharynx), there are a number of other parameters that affect the shape of the vocal tract. Foremost among these are movements of the lips. There are probably only three degrees of freedom involved: the distance between the upper and lower lip (lip height); the distance between the corners of the lips (lip width); and the degree of lip protrusion. In most languages the specifications of lip position in contrasting sounds do not require this number of degrees of freedom. But systematic phonetic differences between languages must also be taken into account. Thus French and German both have front rounded vowels, but there may be less lip protrusion in French.

The degree of velic opening is a well known parameter, and needs no further comment here. Similarly, it is well established that larynx raising and lowering is a controllable gesture that may occur in (among other sounds) different kinds of stop consonants.

There is more disagreement on the parameters required for characterizing glottal states. Despite the elaborate description of what is humanly possible that has been given by Catford (1977), it seems to me that languages use controllable differences in only three parameters: the distance between the arytenoid cartilages (glottal aperture), which is of course, the physiological parametric correlate of oppositions such as voiced-voiceless; the stiffness and mass of the parts of the vocal cords that may vibrate (glottal tension), which may be varied to produce different phonation types such as creaky voice; and the degree of stretching of the vocal cords (glottal length), which correlates most highly with the rate of vibration (the pitch).

The final parameter is lung volume decrement, the prime source of energy for nearly all speech sounds. This is highly correlated with the subglottal pressure, but should not be confused with it. It appears from the work of Ohala (1974) that speakers control the amount of work done by the respiratory system (the rate of decrease of lung volume), rather than the subglottal pressure. Thus they will produce a given amount of power for a given kind of word, irrespective of whether it contains a voiceless aspirate (which will cause a fall in the subglottal pressure) or a glottal stop (which will cause an increase).

Most speech sounds have a unique specification in terms of these 16 parameters. MacNeilage's report may give a slightly wrong impression in this respect. It is not quite correct to say that "Ladefoged et al. (1972) showed that ... there is a considerable variation of tongue configurations adopted by different speakers producing the same vowel." We showed only that different speakers used different degrees of jaw opening to offset different degrees of movement of the tongue relative to the mandible. If by "tongue configurations" one means vocal tract shapes, then one can observe very few differences between speakers.

There are probably only two major ways in which variations in one parameter may lead to no change in the speech sound produced because they are offset by variations in another parameter. The first is the use of larynx lowering to offset decreases in lip rounding (Atal et al. 1977, Riordan 1977). The second is the use of increased respiratory power (lung volume decrement) to offset decreases in the stretching of the vocal cords (glottal length). There may also be variations among the three lip parameters that can be used to compensate for one another. But the data of Atal et al. (1977) on parameterized tongue shapes, and our own similar data, indicate that there are no cases in which a given sound can be produced with the same lip and larynx position, but with two different tongue shapes, as long as the tongue shape is characterized by only two parameters. There are well known cases involving additional parameters, such as American English rhotacized vowels that may be produced in two different ways (Uldall 1958). There may also be variations in pharynx width that can compensate at least in part for variations in front raising and back raising to produce similar tongue shapes in vowels. But apart from the case of rhotacized vowels, I doubt that there are two distinct tongue shapes that produce the same sound.

The 16 parameters listed are hypothesized to be a necessary and sufficient set for linguistic phonetic specifications. Some of them are far from fully defined, but they are all susceptible of precise numerical specification. They are potentially the things that are controlled in speech production. As MacNeilage indicates, we do not yet know whether speech production involves specifying a sequence of targets or whether some form of action theory specification is preferable. The parametric approach outlined above is equally applicable in either case. Very tentatively, Table 1 is offered as a summary of what we use in speech production.

References

Atal, B., J.J. Chang, M.V. Mathews, and J.W. Tukey (1978): "Inversion of articulatory to acoustic transformation by a computer sorting technique", JASA 63, 1535-1555.

Catford, J.C. (1977): Fundamental Problems in Phonetics, Edinburgh: Edinburgh University Press.

Fujimura, O. and Y. Kakita (1979): "Remarks on quantitative description of the lingual articulation", in Frontiers of Speech Communication Research, B. Lindblom and S. Öhman (eds.), London: Academic Press.

Harshman, R., P. Ladefoged, and L. Goldstein (1977): "Factor analysis of tongue shapes", JASA 62, 693-707.

Ladefoged, P., J. DeClerk, M. Lindau, and G.A. Papçun (1972): "An auditory-motor theory of speech production", UCLA Working Papers in Phonetics 22, 48-75.

Ladefoged, P., R. Harshman, L. Goldstein, and D.L. Rice (1978): "Generation of vocal tract shapes from formant frequencies", JASA 64, 1027-1035.

Ladefoged, P. and R. Harshman (1979): "Formant frequencies and movements of the tongue", Frontiers of Speech Communication Research, B. Lindblom and S. Öhman (eds.), London: Academic Press.

Lindblom, B., J. Lubker, and T. Gay (1978): "Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation", JPh (in press).

Lindblom, B.E.F. and J.E.F. Sundberg (1971): "Acoustical consequences of lip, tongue, jaw and larynx movement", JASA 50, 1166-1179.

Lindau, M. (1979): "The feature Expanded", JPh (in press).

Maddieson, I. (personal communication): "Phonological inventories of the languages of the world".

Ohala, J. (1974): "A mathematical model of speech aerodynamics", Proc. Speech Comm. Sem., Stockholm, 65-72, Uppsala: Almqvist and Wiksell.

Ohala, J. (1974a): "Phonetic explanation in phonology", Papers from the parasession on natural phonology, 251-271, Chicago: Chicago Linguistic Society.

Riordan, C.J. (1977): "Control of vocal-tract length in speech", JASA 62, 998-1002.

Uldall, E.T. (1958): "American 'molar' r and 'flapped' r", Revista do Laboratorio de Fonetica Experimental, Coimbra 4, 103-106.