ARIPUC 18, 1984, p. 255-276

SOME BASIC VOWEL FEATURES, THEIR ARTICULATORY CORRELATES AND THEIR EXPLANATORY POWER IN PHONOLOGY*

ELI FISCHER-JØRGENSEN

The classical articulatory vowel features, vowel height (or degree of openness) and front-back, have been criticized as physiologically inexact and in reality based on a misinterpretation of auditory impressions. Some want to replace them by new features, partly based on constriction (Wood), others want to interpret them exclusively in auditory terms (Ladefoged). It is argued in this paper that the classical articulatory vowel features are not as inexact physiologically as maintained by their critics, and that they are indispensable in phonological descriptions and in this respect more useful than e.g. Wood's feature system.

I. INTRODUCTION

For the description and explanation of universal constraints and tendencies in phonological patterns and phonological change one needs detailed models of speech production and speech perception, but it is also necessary to have a general frame of reference in the form of a system of more abstract phonetic dimensions according to which the speech sounds of a language can be grouped into classes. These dimensions which, according to a now generally accepted but not quite unambiguous terminology, are also called features, and which can be considered potentially distinctive, must on one hand have correla-

---

*) *This is a revised and enlarged version of my contribution to the symposium "Phonetic Explanation in Phonology" at the Tenth International Congress of Phonetic Sciences, Utrecht 1983. Particularly the phonetic discussion is more detailed in the present paper.*

tion to speech production and speech perception, on the other be adequate for the description of phonological patterns and rules.

In the present paper I will consider vowel features only, and only some basic articulatory features which have been the subject of debate in recent years, i.e. vowel height (or degree of openness), front-back, and tenseness.

II. CRITICISM OF THE TRADITIONAL SYSTEM

Since the days of Bell and Sweet it has been the tradition to describe vowel systems by means of the basic features high/low, front/back and rounded/unrounded and, for some languages, tense/lax. This system, which was defined in articulatory terms, has been severely criticized for not covering the articulatory facts.

As early as 1910 E.A. Meyer found that lax [ɪ] may have a lower tongue position than generally assumed, e.g. North German [ɪ] is often lower than [e:]. G.O. Russell (1928 and 1936) attacked the traditional system of classical phonetics on the basis of a large number of X-ray photos, particularly of American English, but also of German and French. He found that often the tongue position was not as one should expect; [ɪ] might, e.g., be lower than [e], and for the back vowels there were great discrepancies. For [ɒ], [ɔ] and [ɑ] the point of articulation was rather in the pharynx, whereas [u] might almost be a front vowel. On the whole, the importance of the pharynx cavity for vowel sounds had not been realized. Russell did not have much influence, though, perhaps because of his very aggressive tone. He called it "a wildly unscientific absurdity" to listen to speakers' sounds and then record them in terms of their physiological character, and he characterized the traditional physiological dimensions (particularly high/low) as purely imaginary.

In the sixties the criticism was renewed by Peter Ladefoged (1962 (1967), 1971, 1975, 1976). He maintained, like Russell, that Bell, Sweet and their followers had described their auditory impressions but translated them into physiological terms, and that these did not correspond to a physiological reality. He showed, among other things, that in the X-ray photos of Stephen Jones' cardinal vowels [o] and [ɔ] were lower than front [a], whereas [ɑ] was still lower, as far as the highest point of the tongue is concerned. On the whole, high vowels do not all have the same height ([u] often being lower than [i]) and [u] is often advanced compared to [o]. Ladefoged et al. (1972) also found a rather large individual variation in the articulation of front tense and lax vowels in American English. Ladefoged (1976) finds a much closer agreement between traditional vowel descriptions and acoustic facts. As an example he compares Uldall's placement of the Danish vowels in Jones' cardinal vowel chart and a plot of the formant frequencies of Danish vowels, height corresponding to the frequency of F1, and front-back to the difference between the first and second formant. Joos (1948) had made a similar observation, comparing Jones' cardinal vowel chart with Holbrook and Carmody's X-ray vowel quadrilateral on one hand, and with an acoustic chart of French vowels on the other, and finding a greater similarity in the latter case. - But this is not so astonishing since Jones' cardinal vowels (apart from [i] and [ɑ]) were based on auditory impressions of equal distances.

Nearey (1978), who has analysed the vowels of three American English speakers, also emphasizes that the individual variation is more pronounced in articulation than in the acoustic pattern, and that the traditional vowel features are more closely related to the acoustic aspect.

The most severe attack on the classical vowel system has been made by Sidney Wood (1975a and 1982), based on an analysis of 38 sets of X-ray photos of 15 different languages. He found that [ɪ] and [ʊ] are often lower than [e] and [o], and that the height relations between [o,ɔ] and [a,ɑ] are random. He also emphasizes that the classical system neglects the pharynx cavity completely, and thus cannot account adequately for the relations between vowel production and vowel acoustics, nor is there any clear relation between the dimensions front/back and high/low and the function of the muscles. He concludes that the model is not only inaccurate but irrelevant to the processes of speech production, that it is a complete illusion and must be rejected and replaced by a new model.

III. THE CLASSICAL SYSTEM IS NOT AS INACCURATE AS MAINTAINED BY THE CRITICS

I agree with Catford (1981) that the criticism of the classical system is exaggerated. In the first place, even when accepting the premises of the critics, e.g. that the classical system describes the position of the highest point of the tongue, many X-ray photos show quite a good agreement with the traditional description, as shown for instance by Mona Lindau (1978) for five American English speakers. Ladefoged's own vowels, which he quotes (1976) as an example of bad agreement, are not too bad either. Nobody has ever claimed that all high vowels should have the same height or that front and back vowels of different height should lie each on a straight (vertical) line. For front vowels it is evident that the lower vowels are normally gradually retracted, and the fact that many languages have a somewhat advanced [u] does not invalidate the general classification. Further, Ladefoged's worst example (Stephen Jones' cardinal vowels) (1971) should not be given too much weight. The X-ray photos were taken in 1929, at a time when the technique was much less advanced than today; the subject had a chain lying along the surface of his tongue and another chain through his nose hanging down behind the uvula (Catford

1981), and, moreover, I have been told that he had all the teeth on one side of his mouth extracted in order to get nice pictures! It is, of course, laudable that he suffered so much for science, but it has hardly contributed to the naturalness of his pronunciation.

In the second place, some of the more serious criticisms raised by Ladefoged and Wood lose their force when it is realized that they are mainly objections against Daniel Jones' cardinal vowel chart, which was meant as a practical device for field work and not as a theoretical vowel system, and which differs from the classical system on various points. As mentioned above, it was in the first place an auditory chart, but both Jones himself and his pupils also tended to interpret it physiologically in the sense that it indicated the position of the highest point of the tongue. Ladefoged (1971, p. 67) says that for the past hundred years vowels have been described in terms of the highest point of the tongue. But I do not think this is correct. Catford (1977 and 1981) states that neither Bell nor Sweet used the term. But he is not sure about their followers since Jespersen mentions it as a current term in 1889. But Jespersen does not use the term "highest point of the tongue" here. He talks about the highest <u>part</u> of the tongue. Thus, what he discusses is whether one should talk about tongue height or (as Jespersen and many others preferred) about the distance from the palate. As far as I can see, none of the founders of classical phonetics mention the highest point of the tongue, not, e.g., Jespersen, nor Sievers, nor Viëtor, nor Passy, nor Storm. They therefore often use the terms "degree of openness" or "distance" instead of "height". The same is true of their followers on the European continent with the exception of some more recent works which are influenced by Jones. It seems to be a particular British tradition introduced by Daniel Jones in "The Pronunciation of English" (1909) (not in his "Outline of English Phonetics" 1918, as I said in Fischer-Jørgensen 1983). It is an attempt at a more precise description, but the highest point of the tongue is rather variable and much too precise a concept to be used in a general vowel system. What is implied in the classical system is the distinction between an advanced and a retracted tongue body (sometimes including a central position), and the overall distance between the articulating part (front or back or central) and the palate.

Similarly, the fact that e.g. English [ɪ] and [ʊ] are often lower than [e] and [o] can be used as an argument against Jones' placement of the former vowels above [e] and [o] in the vowel chart, although, as Catford remarks (1981), nothing in the system prevents a placement of a relatively low American English [ɪ] below [e]. What is disturbing is, however, that they will be placed sometimes above and sometimes below [e] and [o] for different speakers. The cause of the trouble is that Jones does not like to recognize tenseness as a separate vowel dimension but prefers to describe [ɪ] as a lowered and retracted [i], and [ʊ] as a lowered and less rounded [u]. If tenseness is recognized as a separate dimension, then [ɪ] and [ʊ] are simply the highest lax vowels, and their exact relation to [e] and [o] is not relevant. This is the position taken by various phoneticians within the classical tradition. The height of [ɪ] and [ʊ] thus cannot be used to criticize the traditional vowel system as such. But the attitude to this dimension has been somewhat hesitating, also within classical phonetics. One reason is probably that Bell's description was evidently wrong. He used the terminology primary vs. wide and thought that the wide (e.g. lax) vowels have a retracted soft palate and an expansion of the pharynx. Sweet, who uses the terminology narrow/wide, explains the difference as depending on the shape of the tongue, which is more convex in narrow (tense) vowels and more flattened and relaxed in wide (lax) vowels. The narrowing is thus not the result of raising the whole body of the tongue with the help of the jaw as in [i] versus [e], but of bunching up the part with which the sound is formed. Thus, passing from [i] to [e] one does not go via [ɪ]. This description is much more to the point, but it was not generally accepted. Jespersen (e.g. 1897-99 and 1914) gave a somewhat different formulation; he described the difference as a thinner vs. broader channel between the articulating part of the tongue and the palate, which did not necessarily involve a generally lower tongue position. Others emphasized the muscle tension as the decisive difference (e.g. Sievers 1901), and the difference is now (therefore) generally called tense-lax. Almost all British phoneticians have followed Jones in not recognizing tense/lax as an independent dimension, whereas a good number of Scandinavian, German, and Dutch phoneticians have accepted the distinction. The difference is also clearer in German and Dutch than in English, where all vowels are relatively lax, but where the difference is clear at least for high vowels. But it is true that many phoneticians found E.A. Meyer's and Russell's findings that [ɪ ʏ ʊ] often had a lower tongue position than [e ø o] disturbing because they found them auditorily closer to [i y u]. However, this auditory impression may be caused partly by the fact that [ɪ ʏ ʊ], e.g. in German, are the highest lax vowels and thus systematically high, and also by orthography; cp. that Danish listeners described German [ɪ ʏ ʊ] - cut out of words and presented in isolation - as being close to Danish [ɛ: œ: ɔ:] (Fischer-Jørgensen 1973 (1975)).

The tense/lax distinction has been widely accepted in modern phonology. The definitions given by Jakobson and Halle (1956) and by Chomsky and Halle (1968) are based on the idea of a difference of muscular tension leading to a greater deviation from the neutral position of the vocal tract. - The terms tense/lax have also been applied to the difference which is the basis of vowel harmony in a number of West African languages. Ladefoged (1964) uses this terminology (though only as a tentative label); but at the same time he shows that the decisive difference, at least in Igbo, lies in an advancement or retraction of the tongue root. Halle and Stevens (1969) think that this is also the main characteristic of the tense/lax difference found in, e.g., the Germanic languages and they therefore propose to use the term "advanced tongue root" in both cases. However, Stewart (1967), Ladefoged et al. (1972) and Lindau et al. (1972) showed that these features should not be confounded (cf. also Lindau 1978, who proposes to call the African feature "expansion"). Whereas the tongue root movement is really an independent feature in various African languages, it is not consistent in English, and it is closely related to tongue height differences. Acoustically the two features are also different (cp. the acoustic graphs of African languages in Lindau et al. 1972, and the graphs of German in Jørgensen 1969). Finally, expansion does not have the close relation to length generally found for tense/lax.

Since the articulatory correlates of the tense/lax difference in English seem to vary according to speakers (Ladefoged et al. 1972), Ladefoged (1975) proposes to define the difference in purely distributional terms (lax vowels being those which do not occur in open syllables), whereas Lindau (1978) gives an acoustic definition (peripheral versus central in the formant chart).

More recently Wood (1975b and 1982) has undertaken an extensive investigation of the articulatory characteristics of tense and lax vowels in a number of languages. He finds that in various languages [i] and [ɪ] have approximately the same jaw opening, whereas for [e] and [ɛ] it is lower, conversely [i] and [e] have approximately the same tongue bunching, whereas [ɪ] and [ɛ] have a flattened tongue. There are exceptions (cp. that the six informants investigated by Ladefoged et al., 1972, all have the distinction in jaw opening described by Wood, but only three have a clear difference in tongue lift). On the whole, however, the tendency is very clear (cp. also Fischer-Jørgensen 1973). It also appears from the measurements of jaw opening for Dutch vowels by Zwaardemaker and Eijkman (1928) and by Kaiser (1941), and, e.g., from X-ray photos of Telugu taken by Nagamana Reddy, which I have had occasion to see. - Thus, whether [ɪ] is lower or higher than [e] is a more accidental consequence of the relative extent of the two movements involved (tongue flattening and jaw opening), but at any rate [ɪ] has less constriction than [i] at the place of articulation, and Wood states that this is true of all other lax vowels compared to their tense counterparts, with the exception of [o/ɔ]. At the same time lax vowels have a narrower pharynx cavity (except for tense [ɑ] versus lax [a]), and less pronounced lip activity. Wood has shown, through synthesis, that these different factors may be of different acoustical importance. It is, however, clear that they can all be regarded as consequences of one single difference: tense versus lax articulation, involving a flattening of the tongue which is closely connected with less space in the pharynx as well as less jaw opening and less pronounced lip activity. There exist few EMG-recordings of the muscular tension involved in the production of tense and lax vowels, and only American English has been investigated. There seems to be a clear difference in the tension of the genioglossus (Smith and Hirano 1968, Raphael and Bell-Berti 1975, Alfonso and Baer 1981, and Alfonso et al. 1982) and of the inferior longitudinal (Raphael and Bell-Berti 1975), whereas there are contradicting results for the styloglossus for [u/ʊ] (EMG-recordings of German and Dutch vowels would be an obvious task for the future).

It thus seems clear that tenseness is an independent vowel dimension (which is often - but not always - combined with length). Height should therefore be indicated separately for tense and lax vowels.

The most serious objection against the traditional height dimension seems to be the irregular height relations between back vowels, as described by Wood. It was, however, also Jones who, for practical reasons, placed the unrounded back vowel [ɑ] in the same column as the rounded vowels [u o ɔ] because, in many languages, [ɑ] is the only low back vowel, whereas the higher back vowels are generally rounded. But from the point of view of a general system of vowel features [ɑ] does not belong in this series but in the series of unrounded back vowels [ɯ ɤ ʌ ɑ]. Thus, the fact that [o] and [ɔ] may be lower is only crucial if height is taken to be an absolute property, not if it is taken to be a relative property within each series of rounded and unrounded, front and back vowels, a view which would be in accordance with Jakobson's conception of distinctive features.

Thus, if tenseness is considered a separate dimension and height is taken to mean the relative distance between the articulating part of the tongue and the palate within each series of rounded or unrounded, tense or lax, front or back vowels, most of the inconsistencies between these traditional labels and the articulatory facts disappear.

What remains of the criticism is that the pharynx cavity, which could not be observed at the time when the classical system was set up, has been neglected. The description does not take account of the fact that the most narrow constriction for [a] and [ɑ] and generally for [o ɔ] is in the pharynx, and as degree and place of the constriction in the total vocal tract are essential for the calculation of the acoustic output, the classical system is not the most adequate starting point for such calculations. It is true that since there are strong constraints on the possible positions of the tongue body (e.g. a low back vowel will necessarily have a narrow pharynx, and a high front vowel a wide pharynx), it has been possible to set up correlations between the features of the classical system and formant frequencies, but the causal relations are not clear if the pharynx is not taken into account.

Wood (1975a) also argues that the connection between vowel articulation and muscle activity becomes much clearer in a model based on place of constriction than in the traditional model. This is not quite so convincing. The relations are complicated in either case (for the relation between muscle activity and the traditional dimensions, see Catford 1977, p. 186 and Halle 1982).

IV. NEW FEATURE SYSTEMS

It is evident that models of vowel production must be based on these more recent insights, i.e., they must take account of the total vocal tract. And it is possible to construct new feature systems, which are connected more closely to these new models.

However, phoneticians do not agree on the number of places of constriction. Catford (1977) tentatively sets up six points, but he does not find this very useful for the description of vowel systems. Lindblom and Sundberg's model of vowel production (1969 and 1971) is based on place and degree of constriction, jaw position, width and height of lip opening, and larynx height. They set up three places of constriction: palatal [i e ɛ], velar [u], and pharyngeal [o a ɑ]. Tongue height is a derived parameter, controlled by means of both jaw position and tongue raising relative to the jaw. In optimal articulation jaw and tongue go together, but there may be compensations. In the 1969 paper a system of binary features is based on this model and applied to the Swedish vowels. It contains the features palatal, pharyngeal, velar, close, open, and labial. It is not, however, mentioned in later publications, and I will therefore not discuss it in detail.

Wood operates with four places of constriction: the hard palate, the soft palate, the upper pharynx, and the lower pharynx. There are thus four categories of vowels: palatals (all front vowels except [æ] and [a]), velars ([u] and [ʊ]), velopharyngeals ([o] and [ɔ]), and low pharyngeals ([æ a ɑ]). It is not a continuum, but four discrete places (he thus denies the existence of central vowels). His arguments are partly empirical (the analysis of X-ray photos from many languages), partly theoretical; these are the places where the spectrum is relatively insensitive to moderate tongue displacements according to Stevens' quantal theory (1972) and Fant's nomograms (1960). He also adduces sound typology (1982, p. 72), pointing to the fact that languages with only three vowel phonemes generally use the three places mentioned by Stevens (/i ɑ u/), and if there are five phonemes the normal type is /i ɛ ɑ ɔ u/. But he does not mention that in languages with four vowel phonemes it is extremely rare to find his four basic vowel types /i u o ɑ/, the most common types being /i u ɑ/ plus /ɛ/ or /ɨ/ (cp., e.g., Crothers 1978). The typological argument is thus very weak.

In Wood's system the tense vowels [i u o ɑ] may have lax counterparts with less constriction [ɪ ʊ ɔ a]. Moreover, the jaw may be close or open. This adds a distinction to palatal vowels: [e] and [ɛ] being the open counterparts to [i] and [ɪ]. The velar vowels are considered to be (redundantly) close, the pharyngeal vowels open. Finally, vowels may be rounded or unrounded.

On this basis he sets up a binary feature system (1982, p. 168):

constriction	palatal				palatovelar		pharyngo-velar		low pharyngeal	
vowel	i	ɪ	e	ɛ	u	ʊ	o	ɔ	ɑ	a
palatal	+	+	+	+	+	+	-	-	-	-
velar	-	-	-	-	+	+	+	+	-	-
pharyngeal	-	-	-	-	-	-	+	+	+	+
open	-	-	+	+	-	-	+	+	+	+
round	-	-	-	-	+	+	+	+	(-)	-
tense	+	-	+	-	+	-	+	-	+	-

Wood does not comment on this table. Evidently it has been a problem to translate a place dimension with four places into a binary feature system. This has the consequence that he has had to call his velar place palatovelar, and his upper pharyngeal place velopharyngeal, which blurs the quantal aspect. It would have been more in keeping with his own description to set up a multivalued dimension with four steps.

It may also be argued that the place of constriction is just as variable as the highest point of the tongue. Front vowels may have their maximal constriction at the alveolar ridge (Straka even calls the French palatals alveopalatals, or simply alveolars (1978)). An [o] may have its maximum constriction at the velum, at the uvula, or in the pharynx, depending on the shape of the individual palate, the position of the velum, etc. Wood is aware of this. 1982, p. 142 he says that "place of maximal constriction" is ambiguous in natural speech. What matters is the direction of the tongue body movement. Even with this precaution one may sometimes get into doubt. In Danish [o:] and [ɔ:] the main direction of the movement seems to be towards the soft palate, although they have at the same time a narrower pharynx compared to [u] (see figure 1)*. Danish [o:] is very close, and it might perhaps be described as an *u*-like vowel. This would only require a step "open" for velar vowels.

V. THE USEFULNESS OF THE TRADITIONAL FEATURE SYSTEM AND WOOD'S SYSTEM IN PHONOLOGICAL DESCRIPTIONS

A feature system based on place of constriction may come closer to a model for vowel production (but it seems rather difficult to translate it into a two- or three-dimensional figure resembling the formant chart). Now, as stated in the introduction, a feature system should also be applicable to phonological descriptions, and Wood expressly emphasizes that his

---

*) See p. 276.

model has more explanatory power in this respect than the traditional system. But for this purpose I find the classical system clearly superior.

Wood only allows for two steps in vowel height (or openness), and for the back vowels the traditional two degrees are reinterpreted as a difference of place. But there are many languages with three degrees of opening, and Danish even has four (/i e ɛ æ/). Wood considers [æ] to be a pharyngeal vowel, but in that case it has to be distinguished by a different feature from /ɑ/, which may be considered a separate phoneme in "Advanced Standard Copenhagen" (see Basbøll 1972). According to his practice for English, Wood must then describe /æ/ as lax. The same must be done for /ɛ œ ɔ/, both in Danish and in languages with similar systems. But there is no justification for describing Danish /ɛ œ ɔ/ as lax (apart from the fact that a certain laxing is normally combined with lowering). Danish has both long and short vowels with, in most cases, almost the same quality. In the series /i e ɛ æ/ there is a gradual lowering of the tongue height (as can be seen in X-ray photos and palatograms) and generally a gradual lowering of the jaw, which appeared in a recording of jaw opening for three Danish informants. With the exception of /e,ɛ/ for one of the informants all the differences were statistically significant (the data will be published later). This speaks against considering any of the vowels as lax. As for Swedish, Lindau (1978) quotes a case of diphthongization in the Swedish Skåne dialect which can only be formulated in a simple way on the basis of four degrees of openness. The English vowel shift also requires more degrees of openness to describe the diphthongization of /iː uː/, the development of /eː oː/ to /iː uː/, of /ɛː ɔː/ to /eː oː/, and of /aː/ via /æː/ to /eː/.

Now this might be partly remedied by adding more steps to the close/open dimension without changing the rest of the system, but in phonological rules /ɑ/ does not go with /e ø o/ as it should according to Wood's feature system because it has the same degree of openness. - It behaves like a lower vowel. There are, e.g., languages where /eː øː oː/ are diphthongized to, e.g., /iə yə uə/, whereas /ɑː/ is not diphthongized, for instance Old High German.

There are also a number of well known universal phonetic tendencies connected with vowel height. Low vowels are longer than high vowels, they are pronounced on a lower pitch and have higher intensity. These differences are gradual, i.e. [a] is longer than [ɛ], [ɛ] than [e], [e] than [i], and similarly for pitch and intensity. It is not a difference between two categories high and low. These differences are generally not perceived, but when Danish speakers are asked to manipulate synthetic vowels in duration adjustment tasks, they make [ɛ] longer than [i], and [ɛː] longer than [iː], and the cross-over value between /i/ and /iː/ is smaller than between /ɛ/ and /ɛː/ (Petersen 1974). These are phonetic differences, but they may turn up in historical developments, so that long high vowels become short vowels (e.g. in Dutch), or short low vowels are lengthened (e.g. English /æ/).

However, it is possible to quote one example where /ɑ o ø e/ go together and behave as one class. In French there are common rules of lengthening and distribution for /ɑ o ø/ and (partly) /e/ vs. /a ɔ œ ɛ/. This induced Oluf Thorsen (see Jensen et al. 1971) to describe /a ɑ/ not as low vowels but as pharyngeal vowels. The common property for /ɑ o ø e/ versus /a ɔ œ ɛ/ was then a more "close timbre" vs. a more "open timbre", later called tense vs. lax. I wonder, however, whether it is correct phonetically to call French /a ɔ œ ɛ/ lax vowels. They do not sound lax to me. Wood (1982, p. 139) found the tongue to be lower relative to the mandible in French /ɛ/ than in [e] in three sets, but the jaw is also more open, and in Straka's X-ray pictures (1950) /ɛ œ ɔ/ have a larger jaw opening than /e ø o/, whereas /a/ has a slightly smaller opening than /ɑ/. In the case of the low pharyngeal vowels it is difficult to judge tenseness on the basis of the constriction, because two conflicting tendencies are at work: (1) a narrower constriction at the point of articulation (i.e. the pharynx) in tense vowels, and (2) a narrower pharynx in the lax vowels. Wood considers [ɑ] generally as tense. This is probably true of German, where it is long, as other tense vowels, but in Dutch it is the front [a] that is long.

Anyhow, the French example can be considered to support Wood. But in all the other cases mentioned we need the traditional height dimension, where [a ɑ] are the lowest vowels.

As for the use of four places of constriction, it makes the formulation of various phonological rules and developments rather complicated. In Germanic umlaut back vowels are changed to front vowels before an [i] in the following syllable (and [ɑ] is also raised). In Finnish and Turkish vowel harmony back vowels are changed to front vowels after front vowels. These rules can only be formulated very clumsily in terms of four places of articulation. The same is true of the allophonic fronting of back vowels in palatal surroundings in Russian.

There are, however, a few cases where the feature pharyngeal for vowels might perhaps give a simpler and more explanatory formulation, i.e. in the cases of assimilation of vowels to pharyngeal or uvular consonants, as found in Greenlandic before /ʁ/ and /q/ (mentioned by Wood). However, as /æ/, which becomes [ɑ] in this position, is considered by Wood to be pharyngeal already, whereas this is not true of /i/ and /u/, the formulation will not be simple. The same reasoning is valid for the ʁ-colouring of Danish vowels (described by Basbøll 1972, and by Basbøll and Kristensen 1975). Perhaps it is just as acceptable to say that vowels may be retracted and lowered before pharyngeal consonants.

This raises the more general problem of the utility of having common features for vowels and consonants, whereby assimilations between vowels and consonants could be formulated in a more explanatory way. Jakobson's features grave/acute and compact/diffuse were not satisfactory in this respect. But instead of using traditional consonantal features for the vowels the goal may also be reached by using the traditional vocalic features for consonants, as it was done by Chomsky and Halle (1968) (see the discussion in Fischer-Jørgensen 1975, p. 230ff) and by Halle (1982). Ladefoged (1971) proposes to specify both consonants and vowels in high/low, front/back and in terms of place of articulation (palatal, velar, pharyngeal). This seems somewhat complicated. But it might be possible to call the vowel front-back dimension palatal/non-palatal, and use non-palatal as a cover term for velar, upper pharyngeal (or uvular), and lower pharyngeal. One might then use the specification in the (rather few) cases where it seems adequate.

The general conclusion of this discussion is that we cannot do without the traditional classical dimensions front-back and high-low.

VI. THE AUDITORY INTERPRETATION OF THE DIMENSIONS HIGH-LOW AND FRONT-BACK

Russell, Ladefoged and Wood agree in the assumption that the dimensions high/low and front/back as used by Bell and Sweet were in reality a translation of their auditory impressions into physiological terms and thus an illusion. Catford rejects this assumption. He says that there is ample internal evidence in the works of Bell and Sweet to show that they were really observing tongue positions. Particularly Sweet was very explicit on this point. He criticizes the German phoneticians for basing their vowel systems on auditory similarity instead of production, and he recommends whispering the vowels in order to better feel the muscular sensations and says that training of the vocal organs is a better way of learning sounds than doing it by ear. If they had built on auditory impressions they would also, as Catford remarks, have placed [y] and [ɯ] between [i] and [u] and not set up rounding as a separate dimension. I think it is true that these old phoneticians, who did not have the possibility of looking at X-ray photos, worked really hard to train their muscular sensations, much more than we do nowadays (this is also true of Jespersen). Catford concludes that the classifications of Bell and Sweet were primarily based on "highly trained phoneticians' perception of proprioceptive and tactile sensations, not upon the misinterpretations of auditory sensations" (1981).

Ladefoged does not, however, draw the same conclusion from his criticism of the traditional system as Wood does. It is true that he also wants to set up a new model of speech production (Harshman, Ladefoged and Goldstein 1977). It is based on a factor analysis of 18 cross sections of the vocal tract for 10 English vowels, ending up with two factors: (1) a forward movement of the root of the tongue together with raising of the front part, approximately from [o] to [i], and (2) raising of the back of the tongue, approximately from [ɑ] to [u]. But he does not want to set up a new feature system on this basis.

He finds that the traditional features have proved to be useful for the description of phonological sound patterns (in 1975 he only gives the supplementary terms palatal and velar for front and back), whereas for the purpose of describing the difference between languages we need a larger number of physiological and acoustic parameters. It must be possible to map the phonological features onto basic phonetic parameters, either physiological or acoustic, but this is not necessarily a one-to-one relation (1980). As for the traditional dimensions, he finds that rounding can be described simply in physiological terms, whereas the physiological correlates to front/back and high/low cannot be used in language description. They have, however, clear auditory and acoustic correlates, height corresponding to the frequency of F1, and front/back to F2-F1 (e.g. 1976). I find it difficult to accept this exclusively acoustic-auditory interpretation.

As for the height dimension it is probably true that it has a somewhat simpler connection with its physical than with its physiological correlates but, as argued above, when height is taken to be a relative dimension within each category of vowels, the correlation to the physiological facts is quite good. It is striking that in experiments intended to bring out the auditory dimensions it has not proved quite easy to factor out a height dimension. Many years ago I tried in various ways (Fischer-Jørgensen 1967). Phonetically naive subjects were asked, e.g., to group vowels according to auditory similarity, with the result that very often front vowels [i e ɛ æ] were put into one group, separated from back vowels, whereas high and low vowels were never sorted out, not even if only six vowels [i y u] and [ɛ œ ɔ] were presented. When asked directly if they did not find that [i y u] belonged together, most declared that [u] does not belong with [i] and [y]. When subjects were asked to group vowels in bright and dark, thin and thick, small and large, etc., or to place them on a scale from dark to bright etc., the same dominating dimension from dark to bright came out in almost all cases, irrespective of the pair of adjectives used. Only when subjects were asked about tight/loose and compact/diffuse did something that looked more like the vertical dimension appear, but [i y u] were designated as compact and [ɛ œ ɔ] as diffuse, in contradistinction to Roman Jakobson's terminology. I think this indicates that the subjects were guided mainly by tactile sensations in this case. When, in some cases, subjects were asked to pay more attention to their articulation, [u] got somewhat closer to [i] and [y]. In more recent experiments, based on more refined methods, when for instance subjects are asked to judge the similarity between sounds presented in triads and the results are factor-analysed, vowel height also turns up (see, e.g., Terbeek 1977). But, altogether, the evidence for "height" being a specifically auditory feature is not very strong.

Moreover, a number of phonetic and phonological rules and developments involving height are better understood when described in articulatory terms. The relatively longer intrinsic duration of lower vowels can hardly be explained from an auditory point of view, whereas it may be assumed that a more extensive jaw movement takes longer time. Preliminary measurements of jaw opening show that the maximum aperture is reached later in open vowels. It may not be economical to accelerate the movement so much that full compensation is achieved. As for intrinsic Fo there are also some plausible articulatory explanations. In the case of palatalization of velars before front vowels, which Ladefoged (1972) also describes as an articulatory development, the difference in palatalizing power between [i e ɛ a] must also be explained on the basis of articulation. But the regularity of F1-correspondences point to an auditory adjustment.

I therefore think that the height feature should be considered to have clear physiological as well as auditory correlates, which both play a role in phonological systems and developments.

As for the dimension front/back it is more complicated. Here two articulatory dimensions, front/back and rounded/unrounded combine to form one auditory dimension: the dark/bright dimension, which was often used in pre-Bell vowel systems. This is, certainly, a dominating auditory dimension, which shows up in experiments with auditory similarity and in phonetic symbolism; it is also prevalent in the patterning of vowel systems (cp. Trubetzkoy's (1939) "Helligkeit" dimension) where /u/ and /i/ are extremely common because they are maximally different in a two-dimensional auditory vowel space, whereas /ɯ/ and /y/, although also distinguished by two articulatory features, are rare (cp. Crothers 1978 and Lindblom 1980), because they are auditorily intermediate between /i/ and /u/. As shown by Ladefoged (1967), even trained phoneticians have difficulty in distinguishing between front rounded and back unrounded vowels. It is also very difficult to elicit rounding as a separate feature in experiments with auditory similarity (cf. Terbeek 1977).

However, in many phonological rules and developments the articulatory dimensions front/back and rounded/unrounded are kept apart. For instance, in Russian there is an allophonic contextually conditioned variation between front [i] and mid [ɨ] and between back [u] and fronted [ü] with preservation of the rounding distinction. Further, in Finnish vowel harmony only the front/back dimension is at work, and in Turkish the front/back harmony and the rounded/unrounded harmony function separately according to different rules. Similarly in i-Umlaut back vowels become front vowels, but the rounding difference is preserved, and conversely, rounding of vowels in labial environment does not involve a change in place of articulation. It is possible that perception may play a role at the last stage of the i-Umlaut, where the [i] of the ending may have become so weak that the listener did not hear it and therefore perceived the front feature of the stem vowel as an independent feature (Ohala 1981), but in its origin it must have been a mainly articulatory process. On the whole, it seems more plausible to explain such processes of assimilation in motor terms as an anticipation of an articulatory position.

Thus I do not think we can do without a front/back dimension defined in articulatory terms. It is probably necessary to operate both with two articulatorily defined dimensions: front/back and rounded/unrounded, and with an auditorily defined dimension of brightness which has a causal relation to a combination of the two articulatory dimensions. They all seem to operate in phonology and thus to be in some sense psychologically real. The articulatory features seem to be at work in assimilatory rules and developments, whereas the auditory feature evidently plays a role in the structure of vowel systems. This is in good agreement with Lindblom's assumption that the necessity for sufficient auditory distance between phonemes is the most important determinant factor in the structure of vowel systems.

However, the relative importance of the horizontal and the vertical dimension has given some problems. In Liljencrants and Lindblom 1972 the possible acoustic vowel space was calculated on the basis of the speech production model set up by Lindblom and Sundberg (1969 and 1971). The calculation was made in terms of formant frequencies. In order to get a two-dimensional space the frequency of formant 2 corrected with respect to formant 3 was chosen as one dimension, and the other dimension was the frequency of formant 1. By transformation into the mel scale an approximation to an auditory space was obtained. By means of a computer program it was calculated where the vowels should be placed in this space if maximal perceptual distance between all vowels should be obtained for different numbers of vowel phonemes (from 3 to 12). The result was compared with known data on actual vowel systems. The prediction turned out to be quite good for vowel systems with 3 to 6 vowels, but above that limit the model generated too many high vowels, i.e., the horizontal dimension was utilized more than the vertical dimension, in contradiction to what is the case in natural vowel systems.

In order to bring the model in better agreement with actual vowel systems it was modified so that the function of the peripheral auditory system was taken more directly into consideration, i.e., instead of formant analysis a filter analysis based on critical bands was used, and masking and non-linear frequency response were taken into account. Moreover, the idea of "maximal" auditory distance was replaced by "sufficient" distance. This model is compared with the older model in Lindblom 1980. It produces a smaller number of high vowels. There are two different versions of the new model. In one of them, which should be closest to the auditory system, phons are transformed into sones, but this version gives less good results for vowel systems with a small number of vowel phonemes, whereas the other version operating on phon/Bark gives better results for systems with few vowels, but somewhat less reduction of the number of high vowels in systems with many vowels. None of the two versions generate the common seven vowel system [i e ε a ɔ o u]. Lindblom suggests that the vertical dimension may play a greater role in actual systems because F1 has high intensity and thus is more resistant to noise.

I should like to suggest a different explanation. Perhaps the auditory distance between [i] and [u] really is felt as relatively long compared to the vertical distance between [i] and [a]. The experiments with Danish subjects mentioned above seem to support this, and the same appears from the dominant use of the dark-bright opposition in sound symbolism. It may be objected that Danish [i e ɛ æ] are closer together than in most other languages. But why then do we keep them apart phonologically even though we feel them as auditorily related? - I suppose production plays a role in this connection. As demonstrated by Lindblom and Sundberg the simplest way to produce differences in vowel height is by raising and lowering the mandible. Now in the first place the proprioceptive sensitivity seems to be more developed for jaw movements than for advancing or retracting the tongue. This may have something to do with the fact that jaw opening and closing is used for other biological purposes, e.g. eating. Moreover, it is visible. (It may happen that a student starting a phonetics course believes that he produces an [e] by retracting his tongue, but he will not maintain that he produces an [a] by closing his mouth.) Finally, steps in jaw movement have a simple one-to-one correlation with steps in F1 and thus with steps in auditory impression of the series [i e ɛ a], whereas the series [i y ɯ u] requires a complicated interplay of tongue and lip movements. Vowel height is a physiologically simpler dimension, and therefore utilized more extensively.

VII. CONCLUDING REMARKS ON EXPLANATION IN PHONOLOGY

Finally I should like to admit that I have used the terms explanation and explanatory in a somewhat slipshod way. But there are many kinds and steps of explanation. When I argue that front-back is more explanatory than four places of articulation in i-Umlaut and vowel harmony rules, the point is that 'back' comprises more cases in one rule, and generalization is a first step in explanation. Distinguishing more than two heights also allows more generalizations, e.g. saying that mid vowels [e ø o] have diphthongized, which is not possible if [ɑ] is considered to have the same height.

But when I argue for an articulatory interpretation of front-back in assimilatory developments, it is because the development can then be described by plausible production mechanisms. Those who, like Roger Lass (1980), require that explanations must be deductive cause-effect explanations which permit prediction would call the explanations mentioned above "understanding" and not "explanation". But that depends on how the word "explanation" is defined. Here it is used in a wider sense. You can never predict a concrete sound change, but you can sometimes explain it afterwards, and you may assume with high probability that if there is a change it will go in a certain direction. There are very strong phonetic constraints on phonological systems and on sound change, and I agree with

John Ohala (e.g. 1983) that it is an important task for phonetics to find these constraints.

ACKNOWLEDGEMENT

I am grateful to Nina Thorsen for improvements of various formulations.

REFERENCES

Alfonso, P. and Baer, Th. 1981: "An electromyographic-cinefluorographic-acoustic study of dynamic vowel production", *Status Rep. Speech Res., Haskins Labs. 65*, p. 109-124

Alfonso, P.J., Honda, K., Baer, T. and Harris, K.H. 1982: "Multichannel study of tongue EMG during vowel production", *Paper presented at the 103rd Meeting of the Acoustical Society of America, Chicago, April 26-30*

Basbøll, H. 1972: "Some conditioning phonological factors for the pronunciation of short vowels in Danish with special reference to syllabification", *Ann. Rep. Inst. Phon. Univ. Cph. 6*, p. 185-210

Basbøll, H. and Kristensen, K. 1975: "Further work on computer testing of a generative phonology of Danish", *Ann. Rep. Inst. Phon. Univ. Cph. 9*, p. 265-291

Catford, I. 1977: *Fundamental Problems in Phonetics* (Edinburgh University Press)

Catford, I. 1981: "Observations on the recent history of vowel classification", *Towards a History of Phonetics* (eds. Asher and Henderson) (Edinburgh University Press), p. 19-32

Chomsky, N. and Halle, M. 1968: *The Sound Pattern of English* (New York: Harper and Row)

Crothers, J. 1978: "Typology and universals of vowel systems", *Universals of Human Language* (eds. Greenberg, Ferguson and Moravcsik), p. 93-152

Fant, G. 1960: *The Acoustic Theory of Speech Production* (The Hague)

Fischer-Jørgensen, E. 1967: "Perceptual dimensions of vowels", *To Honor Roman Jakobson*, p. 667-671, Janua Linguarum, Series Major XXXI

Fischer-Jørgensen, E. 1973: "Perception of German and Danish vowels with special reference to the German lax vowels /I Y U/", *Ann. Rep. Inst. Phon. Univ. Cph. 7*, p. 143-194, and (a shorter version) in *Auditory Analysis and Perception of Speech* (eds. Fant and Tatham) 1975, p. 153-176

Fischer-Jørgensen, E. 1975: *Trends in Phonological Theory*, (Copenhagen: Akademisk forlag)

Fischer-Jørgensen, E. 1983: "Vowel features and their explanatory power in phonology", *Abstracts X Int. Congr. Phon. Sc.*, p. 259-265

Halle, M. 1982: "On distinctive features and their articulatory implementation", to be published in: *Natural Language and Linguistic Theory*, 1983

Halle, M. and Stevens K. 1969: "On the feature "advanced tongue root", *MIT Quart. Prog. Rep. 94*, p. 209-215

Harshman, R., Ladefoged, P. and Goldstein, L. 1977: "Factor analysis of tongue shape", *J. Acoust. Soc. Am. 62*, p. 693-707

Jakobson, R. and Halle, M. 1956: *Fundamentals of Language*, (The Hague: Mouton)

Jensen, O. Kongsdal, Landschultz, K. and Thorsen, O. 1971: *Fransk Fonetik* (University of Copenhagen)

Jespersen, O. 1889: *The Articulation of Speech Sounds Represented by Means of Analphabetic Symbols* (Marburg: Elwert)

Jespersen, O. 1897-1899: *Fonetik* (Copenhagen: Det Schubotheske Forlag)

Jespersen, O. 1914: *Lehrbuch der Phonetik* (Leipzig: Teubner)

Jones, D. 1909: *The Pronunciation of English* (Cambridge University Press)

Jones, D. 1918: *An Outline of English Phonetics* (Cambridge: Heffers and Sons)

Joos, M. 1948: *Acoustic Phonetics*, *Language Monographs 23*

Jørgensen, H.P. 1969: "Die gespannten und ungespannten Vokale in der norddeutschen Hochsprache mit einer spezifischen Untersuchung der Struktur ihrer Formantfrequenzen", *Phonetica 19*, p. 217-245

Kaiser, L. 1941: "Biological and statistical research concerning the speech of 216 Dutch students, III", *Archives néerlandaises de phonétique expérimentale XVII*, p. 92-118

Ladefoged, P. 1962: *The Nature of Vowel Quality*, *Revista do Laboratório de fonética experimental da faculdade de letras da universidade de Coimbra*, and *Three Areas of Experimental Phonetics*, 1967, p. 50-142

Ladefoged, P. 1964: *A Phonetic Study of West African Languages West African Language Monographs 1*, (Cambridge University Press)

Ladefoged, P. 1971: *Preliminaries to Linguistic Phonetics*, (University of Chicago Press)

Ladefoged, P. 1972: "Phonological features and their phonetic correlates", *Journal of the IPA 2*, p. 2-12

Ladefoged, P. 1975: *A Course in Phonetics* (New York: Harcourt Brace & Jovanovich)

Ladefoged, P. 1976: "The phonetic specification of the languages of the world", *Working Papers in Phonetics, UCLA 31*, p. 3-21

Ladefoged, P. 1980: "What are linguistic sounds made of?", *Language 55*, p. 485-502, also *Working Papers in Phonetics, UCLA 45*, 1979, p. 1-24

Ladefoged, P. 1983: "The limits of biological explanations in phonetics", *Abstracts X Int. Congr. Phon. Sc.*, p. 31-37

Ladefoged, P., De Clerk, J., Lindau, M., Papçun, G. 1972: "An auditory-motor theory of speech production", *Working Papers in Phonetics, UCLA 22*, p. 48-75

Ladefoged, P., Harshman, P., Goldstein, P. and Rice, L. 1977: "Vowel articulations and formant frequencies", *Working Papers in Phonetics, UCLA 38*, p. 16-40

Lass, R. 1980: *On explaining Language Change* (Cambridge University Press)

Liljencrants, J. and Lindblom, B. 1972: "Numerical simulation of vowel quality systems: the role of perceptual contrast", *Language 48*, p. 839-862

Lindau, M. 1978: "Vowel features", *Language 54*, p. 541-563, also *Working Papers in Phonetics, UCLA 38*, p. 49-81

Lindau, M., Jacobsen, L. and Ladefoged, P. 1972: "The feature advanced tongue root", *Working Papers in Phonetics, UCLA 22*, p. 76-94

Lindblom, B. 1980: "Phonetic universals in vowel systems", ms. to be published in *Experimental Phonology* (ed. J. Ohala) (Academic Press)

Lindblom, B. and Sundberg, J. 1969: "A quantitative model of vowel production and the distinctive features of Swedish vowels", *Speech Transm. Lab., Quart. Prog. and Status Rep. Royal Inst. Techn., Stockholm 1/1969*, p. 14-32

Lindblom, B. and Sundberg, J. 1971: "Acoustical consequences of lip, tongue and larynx movement", *J. Acoust. Soc. Am. 50*, p. 1166-1179

Meyer, E.A. 1910: "Untersuchungen über Lautbildung", *Die neueren Sprachen 18 (= Festschrift Viëtor)*, p. 166-248

Nearey, T.M. 1978: *Phonetic feature systems of vowels. Indiana Linguistics Club*

Ohala, J. 1981: "The listener as a source of sound change", *Papers from the Parasession on Language and Behaviour* (eds. Miller et al., Chicago)

Ohala, J. 1983: "The origin of sound patterns in vocal tract constraints", *The Production of Speech* (ed. P.F. MacNeilage) (New York: Springer), p. 189-216

Petersen, N. Reinholt 1974: "The influence of tongue height on the perception of vowel duration in Danish", *Ann. Rep. Inst. Phon. Univ. Cph. 8*, p. 1-10

Raphael, L.J. and Bell-Berti, F. 1975: "Tongue musculature and the feature of tension in English vowels", *Phonetica 32*, p. 61-63

Russell, G.O. 1928: *The Vowel* (Columbus, Ohio State University Press)

Russell, G.O. 1936: "Synchronized X-ray, oscillograph, sound and movie experiments showing the fallacy of the vowel triangle and open-closed theories", *Proc. II Int. Cong. Phon. Sc.*, p. 198-205

Sievers, E. 1901: *Grundzüge der Phonetik* (Leipzig: Breitkopf und Härtel)

Smith, T. and Hirano, M. 1968: "Experimental investigation of the muscular control of the tongue in speech", *Working Papers in Phonetics, UCLA 10*, p. 145-155

Stevens, K.N. 1972: "The quantal nature of speech: evidence from articulatory-acoustic data", *Human Communication, a Unified View* (eds. David and Denes)

Stewart, J.M. 1969: "Tongue root position in Akan vowel harmony", *Phonetica 16*, p. 185-204

Straka, G. 1950: "Système des voyelles du français moderne", *Bull. de la Faculté des Lettres de Strasbourg 28*, p. 172-180, 220-233, 275-284, 268-375

Straka, G. 1978: "Apropos du classement articulatoire des voyelles", *Hamburger phonetische Beiträge 25*, p. 432-460

Terbeek, D. 1977: *A Cross-Language Multidimensional Scaling Study of Vowel Perception, Working Papers in Phonetics, UCLA 37.*

Trubetzkoy, N.S. 1939: *Grundzüge der Phonologie, Travaux du Cercle Linguistique de Prague 7*

Wood, S. 1975a: "The weakness of the tongue arching model of articulation", *Working Papers, Dept. Ling. Lund 11*, p. 55-108

Wood, S. 1975b: "Tense and lax vowels - degree of constriction or pharyngeal volume?", *Working Papers, Dept. Ling. Lund 11*, p. 110-134 and *23*, 1982, p. 157-182

Wood, S. 1979: "A radiographic analysis of constriction location for vowels", *J. Phonetics 7*, p. 25-43, and *Working Papers, Dept. Ling. Lund 23*, 1982, p. 57-75

Wood, S. 1982: *X-ray and Model Studies of Vowel Articulation, Working Papers, Dept. Ling. Lund 23.*

Zwaardemaker, H. and Eijkman, L.P.H. 1928: *Leerboek der Phonetiek* (Haarlem: De Erven F. Bohn).

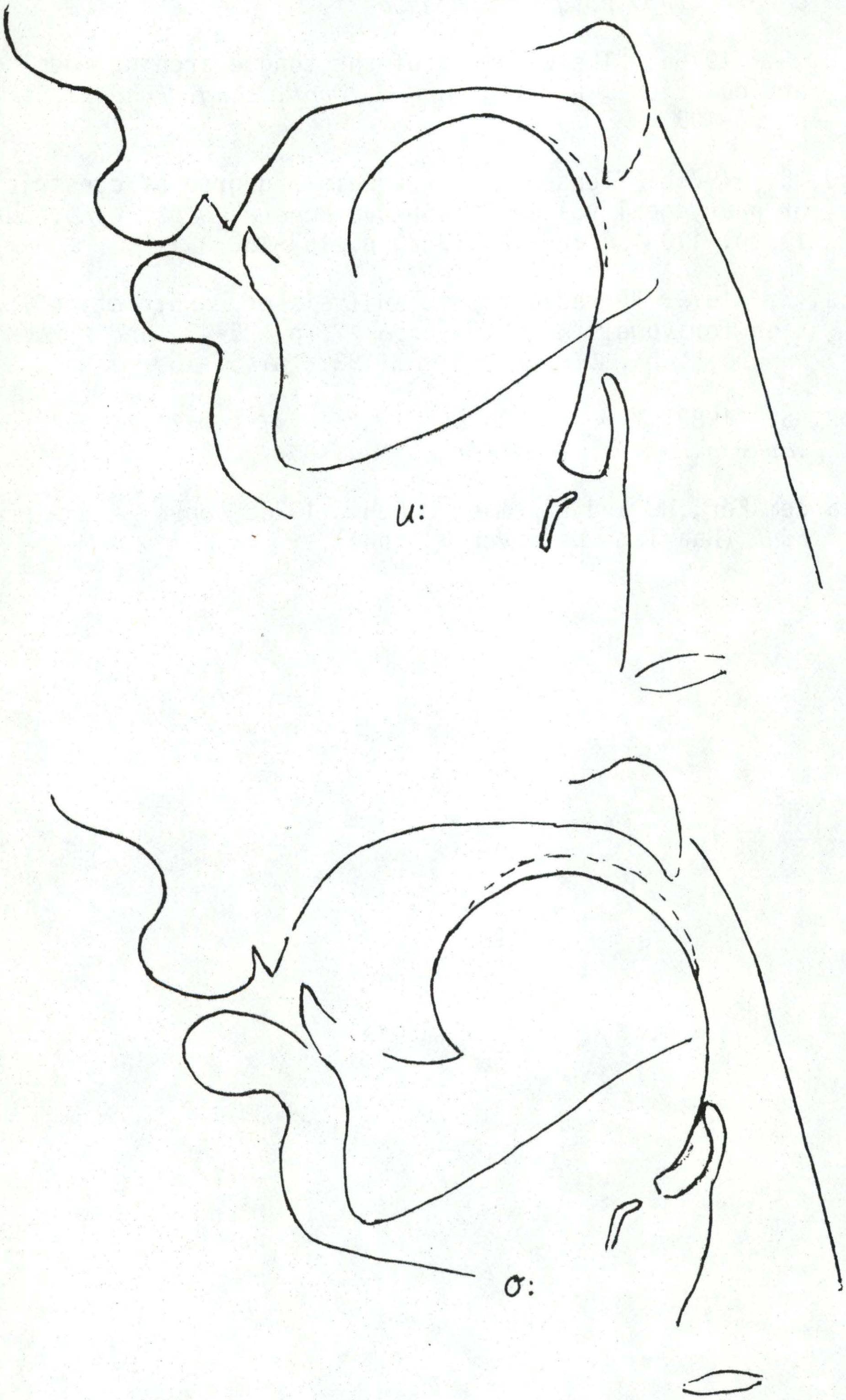

Figure 1

Tracings from X-ray photos of the Danish long vowels /u:/ and /o:/ pronounced by subject KS